

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:37:09 ; Search time 38 Seconds  
(Without alignments)  
4948.535 Million cell updates/sec

Title: US-10-021-571-4

Perfect score: 8864  
Sequence: 1 MSVLISOSVINYVEENIPA.....LHAASSESTGFGERESTIL 1715

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 424699 seqs, 10964683 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCIT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3050.5	34.3	705	US-10-106-698-6378	Sequence 6378, Ap
2	2851.5	32.1	551	US-09-835-788A-17	Sequence 17, Appl
3	2638	29.7	513	US-10-149-819-9	Sequence 9, Appl
4	553	6.2	109	US-09-986-480-371	Sequence 371, App
5	519	5.8	1724	US-09-964-899-43	Sequence 43, Appl
6	437	4.9	740	US-09-835-788A-12	Sequence 12, Appl
7	408.5	4.6	426	US-09-908-711-70	Sequence 70, Appl
8	376.5	4.2	285	US-09-835-788A-18	Sequence 18, Appl
9	376.5	4.2	1327	US-09-972-115A-8	Sequence 8, Appl
10	376.5	4.2	1327	US-09-841-835-2	Sequence 2, Appl
11	370.5	4.2	949	US-09-841-835-10	Sequence 10, Appl
12	370	4.2	1166	US-09-972-115A-6	Sequence 6, Appl
13	370	4.2	1166	US-10-163-587A-15	Sequence 15, Appl
14	357	4.0	1074	US-09-509-196A-2	Sequence 2, Appl
15	348.5	3.9	784	US-10-164-080-7	Sequence 7, Appl
16	347	3.9	802	US-09-964-899-41	Sequence 41, Appl
17	346	3.9	251	US-09-835-788A-13	Sequence 13, Appl
18	345.5	3.9	1333	US-09-972-115A-2	Sequence 2, Appl
19	344	3.9	786	US-10-164-080-2	Sequence 2, Appl

20	344	3.9	786	US-10-299-327-2	Sequence 2, Appl
21	344	3.9	787	US-09-866-050A-334	Sequence 334, App
22	341.5	3.8	679	US-10-339-936-2	Sequence 2, Appl
23	339.5	3.8	1267	US-09-972-115A-4	Sequence 4, Appl
24	320	3.6	835	US-09-947-199-2	Sequence 2, Appl
25	309.5	3.5	673	US-09-841-835-8	Sequence 8, Appl
26	293	3.3	835	US-09-947-199-8	Sequence 19, Appl
27	288.5	3.2	306	US-09-835-788A-19	Sequence 19, Appl
28	285	3.2	1719	US-10-012-896-378	Sequence 378, App
29	285	3.2	1719	US-09-895-793-378	Sequence 378, App
30	285	3.2	1719	US-09-895-814-378	Sequence 378, App
31	285	3.2	1719	US-10-010-940-378	Sequence 378, App
32	285	3.2	1719	US-09-759-143-378	Sequence 378, App
33	285	3.2	1719	US-09-780-663-378	Sequence 378, App
34	285	3.2	1719	US-09-822-827-378	Sequence 378, App
35	265.5	3.0	599	US-09-735-368-2	Sequence 2, Appl
36	265.5	2.9	622	US-10-197-666A-22	Sequence 22, Appl
37	238.5	2.7	622	US-10-197-666A-6	Sequence 6, Appl
38	235	2.6	329	US-09-880-192-62	Sequence 62, Appl
39	235	2.6	329	US-09-758-593A-1	Sequence 1, Appl
40	235	2.6	1054	US-09-798-042-87	Sequence 87, Appl
41	234	2.6	1349	US-10-076-622-573	Sequence 573, App
42	234	2.6	1349	US-10-007-805-573	Sequence 573, App
43	232	2.6	399	US-10-076-622-572	Sequence 572, App
44	232	2.6	399	US-10-007-805-572	Sequence 572, App
45	231.5	2.6	328	US-09-758-593A-11	Sequence 11, Appl

## ALIGNMENTS

RESULT 1  
US-10-106-698-6378  
Sequence 6378, Application US/10106698  
Publication No. US20030109690A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept  
FILE REFERENCE: PA005P1  
CURRENT FILING DATE: 2002-03-27  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: US/10/106, 698  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: PCIT/US00/26524  
PRIOR FILING DATE: 1999-09-29  
PRIOR APPLICATION NUMBER: US 60/157, 137  
PRIOR FILING DATE: 1999-11-03  
NUMBER OF SEQ ID NOS: 8564  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 6378  
LENGTH: 705  
TYPE: PR  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (244)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: MISC FEATURE  
LOCATION: (337)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-6378

Query Match 34.3%; Score 3050.5; DB 9; Length 705;  
Best Local Similarity 98.8%; Pred. No. 3.1e-176;  
Matches 587; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY :|||||  
DB 28 SMCNEDLNWTLALSASKEGVHVEELKCGVLEHRDGMGTALMAGCKRTDVE 121  
122 LLLSHGANPSVGLVYVPIIIMAGRGHADIIVHLLDNGAKVNCSDKXGTPPLVMAARK 181  
88 LLLSHGANPSVGLVYVPIIIMAGRGHADIIVHLLDNGAKVNCSDKXGTPPLVMAARK 146

182 GHECYKHLAMGADVDEGANSMTALIVAVGCTTQSVKELKRNPNVNLTDKGNAL 241  
147 GHECYKHLAMGADVDEGANSMTALIVAVGCTTQSVKELKRNPNVNLTDKGNAL 206  
242 MIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRGHEIYRALLQKADIDINGO 301  
207 MIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRGHEIYRALLQKADIDINGO 266  
302 DKKTALYAVVEKGNATMVDILQCNPDTEICTKDETPLIKATKRNIEVEVLLDKGAK 361  
267 DKKTALYAVVEKGNATMVDILQCNPDTEICTKDETPLIKATKRNIEVEVLLDKGAK 326  
362 VSAVKKGDTPLHAIIRGSRKLAELLRNPKDGLYLRPNKAGTTPNIDCSHOKSLIT 421  
327 VSAVKKGDTPLHAIIRGSRKLAELLRNPKDGLYLRPNKAGTTPNIDCSHOKSLIT 386  
422 QIFGARHLSPETDGMIGYDLYSSALADILSEPTMOPPICVGLYQWGSCKSFLLKLE 481  
387 QIFGARHLSPETDGMIGYDLYSSALADILSEPTMOPPICVGLYQWGSCKSFLLKLE 446  
482 DEMKTFAGQOIEPLFQFSMLIVFTLLCGGLGFATVHPNIGIANSLSFLALLYTF 541  
447 DEMKTFAGQOIEPLFQFSMLIVFTLLCGGLGFATVHPNIGIANSLSFLALLYTF 506  
542 IYIVGGRREGESNMAMVYLSTRLARHIGYLELLKLMFVNPPELPEOTTKALPVRF 601  
507 IYIVGGRREGESNMAMVYLSTRLARHIGYLELLKLMFVNPPELPEOTTKALPVRF 566  
602 DYNRLSSVGETSLAEMTATLSDACEREFGLATRLPVEKTEDTQKKKKKT 655  
567 DYNRLSSVGETSLAEMTATLSDACEREFGLATRLPVEKTEDTQKKKKKT 620

## RESULT 2

US-09-835-788A-17  
; Sequence 17, Application US/09835788A  
; Patent No. US20020077458A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides,  
; FILE REFERENCE: PTO18P1  
; CURRENT APPLICATION NUMBER: US/09/835,788A  
; PRIOR FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: PCT/US00/28666  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/159,585  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 60/167,246  
; PRIOR FILING DATE: 1999-11-24  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 551  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-835-788A-17

Query Match 32.1%; Score 2851.5; DB 10; Length 551;  
Best Local Similarity 99.6%; Pred. No. 2.3e-164;  
Matches 550; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

102 MGGTALMAMCYGRDVVLLSHGANSVYGLQTSYPIIWAAGRGADIVHLLONG 161  
1 MGGTALMAMCYGRDVVLLSHGANSVYGLQTSYPIIWAAGRGADIVHLLONG 59  
162 AKVNCSDKGTPLVMAARKGHECYKHLAMGADVDEGANSMTALIVAVGCTTQSVK 221  
60 AKVNCSDKGTPLVMAARKGHECYKHLAMGADVDEGANSMTALIVAVGCTTQSVK 119  
222 EILKRNPNVNLTDKGNATLMIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRG 281

120 EILKRNPNVNLTDKGNATLMIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVRG 179  
282 HVEIYRALLQKADIDIRGDNKTALYAVVEKGNATMVDILQCNPDTEICTKDETPLIK 341  
180 HVEIYRALLQKADIDIRGDNKTALYAVVEKGNATMVDILQCNPDTEICTKDETPLIK 239  
342 KATKRNIEVEVLLDKGAKVAVDKKDTPLHAIIRGSRKLAELLRNPKDGLYLRPN 401  
240 KATKRNIEVEVLLDKGAKVAVDKKDTPLHAIIRGSRKLAELLRNPKDGLYLRPN 299  
402 NKAGETTPNIDCSHOKSLITQIFGARHLSPETDGMIGYDLYSSALADILSEPTMOP 461  
300 NKAGETTPNIDCSHOKSLITQIFGARHLSPETDGMIGYDLYSSALADILSEPTMOP 359  
462 CVGLYQWGSCKSFLLKLEDEMKTFAQOIEPLFQFSMLIVFTLLCGGLGFATVHPN 521  
360 CVGLYQWGSCKSFLLKLEDEMKTFAQOIEPLFQFSMLIVFTLLCGGLGFATVHPN 419  
522 HPNIGIANSLSFLALLYTFIYVFGGRREGESNMAMVYLSTRLARHIGYLELLKLMF 581  
420 HPNIGIANSLSFLALLYTFIYVFGGRREGESNMAMVYLSTRLARHIGYLELLKLMF 479  
582 NPPELPEOTTKALPVRFITDYNRLSSVGETSLAEMTATLSDACEREFGLATRLPVE 641  
480 NPPELPEOTTKALPVRFITDYNRLSSVGETSLAEMTATLSDACEREFGLATRLPVE 539  
642 KTEDTQKKKKKT 653  
540 KTEDTQKKKKKT 551

## RESULT 3

US-10-149-819-9  
; Sequence 9, Application US/10149819  
; Publication No. US20030044913A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: YUE, Henry  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LU, Dying Alina M.  
; APPLICANT: SHAH, Purvi  
; APPLICANT: IAL, Preeti  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: BUREFORD, Neil  
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES  
; FILE REFERENCE: PF-0760 PCT  
; CURRENT APPLICATION NUMBER: US/10/149,819  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354  
; PRIOR FILING DATE: 1998-12-10; 1999-12-16  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20030044913A1 2215706CD1  
US-10-149-819-9

Query Match 29.7%; Score 2638; DB 9; Length 513;  
Best Local Similarity 99.8%; Pred. No. 1.7e-151;  
Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1203 MNMNGDWHLFRSTYLENRNASHVPEPFLSESSGAPAPHGPARRASHNELPTEL 1262  
1 MNMNGDWHLFRSTYLENRNASHVPEPFLSESSGAPAPHGPARRASHNELPTEL 60  
1263 SSQPTTILNFSPEELNIGLDEGAPRHSNLSWSQOTRTPSLSLNSQDSSIEISLTK 1322



```

Db      61 SSGPTPTLNSEFELTLTGIDECAPRHSNLSMOSQRRRPPSLSLNSQSSSTIEISLTIDK 120
QY      1323 VQAEYDAYREYIAQMSQLEGFGSTTIGRSSPHSTYVMGSSGSGSIHNSLEDEKGD 1382
Db      121 VQAEYDAYREYIAQMSQLEGFGSTTIGRSSPHSTYVMGSSGSGSIHNSLEDEKGD 180
QY      1383 SEKPPDGRKSFIMKRGDVIYSSSGVSTNDASPLDPIEEDKSOOSKILPGKSSSE 1442
Db      181 SEKPPDGRKSFIMKRGDVIYSSSGVSTNDASPLDPIEEDKSOOSKILPGKSSSE 240
QY      1443 RSSLPQTDKLKSGSLRYOKLPSEDESGTEESDNTPLDKDKRAEKKVEYVPSPH 1502
Db      241 RSSLPQTDKLKSGSLRYOKLPSEDESGTEESDNTPLDKDKRAEKKVEYVPSPH 300
QY      1503 SAEPITFTAKKEYLSDALLDKKSSDGVRSSESPNHLHNEVADDSHLEKANLIELE 1562
Db      301 SAEPITFTAKKEYLSDALLDKKSSDGVRSSESPNHLHNEVADDSHLEKANLIELE 360
QY      1563 DSHSGKRGIPHSLSGLDPIITARMGICSEDKKSPSECSLIASSPENMPACOKAYNLNR 1622
Db      361 DSHSGKRGIPHSLSGLDPIITARMGICSEDKKSPSECSLIASSPENMPACOKAYNLNR 420
QY      1623 TPSTVTLNNSAPANRANONFDEMEGIRETSQYILRPSSSPNPTTQNTENLKSMTHRSQ 1682
Db      421 TPSTVTLNNSAPANRANONFDEMEGIRETSQYILRPSSSPNPTTQNTENLKSMTHRSQ 480
QY      1683 RSSYTRLSKDPPELHAAASSESTGFGBERESIL 1715
Db      481 RSSYTRLSKDPPELHAAASSESTGFGBERESIL 513

```

# RESULT 4

```

US-09-986-480-371
; Sequence 371, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: P5500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 371
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-480-371

```

```

Query Match      6.2%; Score 553; DB 9; Length 109;
Best Local Similarity 100.0%; Pred. No. 2e-26;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      302 DNKTALYWAVEKGNATVNRDILQCNPDTEICTGDETPLIKATMRNIYEVELLDGAK 361
Db      1 DNKTALYWAVEKGNATVNRDILQCNPDTEICTGDETPLIKATMRNIYEVELLDGAK 60
QY      362 VSAADKRGDPLHTAIGRSRKLAEILLRNPGRILLYRPNKAGETP 408
Db      61 VSAADKRGDPLHTAIGRSRKLAEILLRNPGRILLYRPNKAGETP 107

```

```

RESULT 5
US-09-964-899-43
; Sequence 43, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.

```

```

; TITLE OF INVENTION: Identification of Genes Involved in
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1724
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-43

```

```

Query Match      5.8%; Score 519; DB 9; Length 1724;
Best Local Similarity 20.5%; Pred. No. 1.3e-22;
Matches 383; Conservative 251; Mismatches 663; Indels 568; Gaps 69;

```

```

QY      14 EENIPALKALLECKDNDEREGCOTPLMTAEGNLEIYKELIKNGANCNLEDNDWT 73
Db      45 KEGHVEVSELLQREANVDAATKKGNTALHLSLAGOAEVYKVTNAGNAOSQNGFT 104
QY      74 ALISAKKEGHVHIVELELLKCGVNLBHRMGWTALMMCYGRDVELELL 124
Db      105 PLYMAQENHLEVVAFLLDNGASQSLATEDEGFTPLAVLQOCHDOVSLLENDTKYVR 164
QY      125 -----SHGANSPTGL----- 135
Db      165 LPLHIAARKDPTKAALLLQNDNNADYESKSGFTPLHIAHYGINVATLLNRAAVD 224
QY      136 ---QYSVPITIAAGRHADIVHLLONGAKVNCSDKGTTPLYAAKRG----- 182
Db      225 PTAARDIPPLHVAASKRGNANMVKLLIDRQAKIDATRGDTPPLHCGASGHEQVVEMLLD 284
QY      183 -----HLECYHLLAMGADVOEGANSMTALIYAVKGYTQS 219
Db      285 RAAPILSKTKNGLSPLHMAATOGDHLNCOVLLQHNVPDDVDVNDVLTPLHVAHAGHYKV 344
QY      220 VREILKRNPNVNLTRKDGNTALMIAKSGHTIEIVODLDAGTYVNIIPRSGDYLIGAVR 279
Db      345 AVLLDOKKANPNAKALNGFTPLHIAKKNRKYVEMELLKHGASIQAVTERGETALHMAAR 404
QY      280 GGHVEIVRALLOKYADIDIRGODNKTALYWAVEKGNATVNRDILQCNPDTEICTGDETP 339
Db      405 SGQAEVRYLVODGAQVEAKADDPPLHISARLGRADIVQOLLQOGASPNAAATTSGYTP 464
QY      340 LTKATMRNIEVEVELLDGAKVSAVDKGDPLHTAIGRSRKLAEILLRNPGRILLY 399
Db      465 LHSAREGHEDVAAPFLDHGASLSTTKSGLTPLHVAHAYDNOKYALLL---DQASPH 521
QY      400 RPNKAGETPNIDC-SHOKSILTOI--FGAR-----HLSPETDGMGLGYDY 444
Db      522 AAANKGYTPLHIAAKKNQMDIATLLEYGADANAATVROGIASVHLAAEGHNDVSLLLG 581
QY      445 SSALDIIEPTMOPICVGLYAQMSGSGSFLKLEDD-----EKTARAOOIE----- 493
Db      582 RNANVNLMSKSGTLP---LHLAAQ-----EDVNVAAEVLVNOGAVNDQOTYV 626
QY      494 ---PLFOFSWILVEITLLCGG-----LGLFAFYVNLGIVASLFLAL 536
Db      627 GPPLPGRKCVHLYT---ANGTPLHQAQOQGHTHINVLNNSPN-ELVYVTEKH 682
QY      537 LYIEFTVIVYFGRREGESNNMAMVLSRILARIIGYLELLIKIMEVNP-----P 584
Db      683 MNV-----PETMNEVLDMSDDEVKRNANPEMLSDXEYISDVEGNRCWTYKIP 730
QY      585 ELPEQTKALPVKFLFTVDNRLSSVCGEISLAEMLATLSDAERREGFLATLFFVFTKE 644
Db      731 KVOEFTVKDITFRKREAFVGLSTLSAGEDAM-----TG 763

```

QY 645 DTG--GKKMK--TCLPSPVIFLIIIGCITIGTILLAFRVDKHLTVNAVLISIAS 699  
 764 DFDKYLGPDLKELGDDSLPABGYMGFSLGA--RSARLVSFVMAVARGSMG-----SR 816  
 QY 700 VVGLAVL--NCRWVOYVLDLSNQRKRLN-----AASKLKLKSEG----- 741  
 817 HNGMRLIIPRCKTAPRTITCRLY--KRHKLANPPMVBEGIASLRYEMGPAQOFLCP 874  
 QY 742 -----FMKVLCEVELMARAKTIDSTFONO-----TRLVIIIDJLADAGQDVLQML 789  
 875 VIVEIPHFGSMGKEHELIVLSNGETKEHOFOSKNDLIELLGMQD-----EEL 926  
 QY 790 DTVRVLFSK-----GPEIAFASD----- 808  
 927 DSPEELGKKRICRITIKDPQYFAVYSRIKESNOGPGGILSTTVLYOASPEGAL 986  
 QY 809 -----PHIILAIKONLSVLRDSNINHDMYRIVLPFL-----NSR 848  
 987 TKRIRVGLAQAPVDEIVAKI--LGNKATFSPITVVEPRRKRKRPITMTIPVPPSGE 1043  
 QY 849 GLSNARK-----FLVTSATNGDVPK--SDTTGIOEDADRVRSONSLCEMTKLGSKTAL 899  
 1044 GYSNGYKGDTPNLRLCISITGTSBQMEDITG--TTPFLFIDCVSFTTNVSAR--- 1097  
 QY 900 NRDYTRRROMQRTITROMSFDLTCLVTEDMFSDISPTQMRLLNIVSVTGRLLRANOI 959  
 1098 -----YGNKGFQKAV-----LEKGPVYDCYGNLAPLTK-----GGOQLVFNFY 1136  
 QY 960 SFNMDRLASMINL--TEOWPYRTSMJLLEETEGIPD----- 995  
 1137 SKRENLPISIKIRTSOEPGRSLFKPKTKTKGIPQTAONLITLPAHKKIEKTRR 1196  
 QY 996 ---QMTLTYERISKNIPTTKDVEPLL--EIDDIRNFEVFLSR-----TPVLVARD 1044  
 1197 OSFASIALKRYSLTEPGMKKKMOSELSDEESTSNSTLSERSRGGQSVTKKARD 1256  
 QY 1045 VAVFLPCTVNDPKLEITADYARAREQI-----STGG 1077  
 1257 KYTE--AAPLKSKEKAGESEKSSRRADALTSLVTRKINRIDITYLEGPIFDYGNISG 1313  
 QY 1078 L-----AYP-----PLPLEHGPAPRAGSYSPSCSSSFNGPAG 1114  
 1314 TSPFADENNVFHDVGYPSLQVELETPTGLHYTP--PTPQODDYSDISIESPL-- 1368  
 QY 1115 GYVSPQPHSSYISGTMGPQHPPYRNGSGPAPGVVLLNSLVANDAVCEKIKOIEGLDOSML 1174  
 1369 -----RTBSRLSDGLVPSQGNIEHSADGP--PVYTAE--DASLEDSKLEDSVPLEM 1416  
 QY 1175 POYCTTIRKANINGVLAOCNIDELKEMNMFQWMLFRSTVLEMRNAESHVVPEDRF 1234  
 1417 PE-----AVDDESQLE--NVCLSEYPOYLQ--NLASGPRDVKRAEPRK 1456  
 QY 1235 LSESSGAPRHPGEPARRASHNELPHTLELSQTPYTLNFSFELNLTGLDEGAPRRHSNSW 1294  
 1457 LGVSE--QOEKKSQSPDEEMMEKIKLS-----LFEDIQ--LEEG-----YES 1495  
 QY 1295 OSQTRTPSLSLNSQSSIEISKLTDRVQAEYRDVREYLAQMSQLEGGGSTTIGSR 1353  
 1496 EEMTEEKYQAILKRVQOAELEMSITG--WONETSSGNLESCAQARVRVG-----LIDRL 1549  
 QY 1354 -SSPH-----STYMGSSSGSGSIHSLNLEOKGKDSPPKDDGKSKSLMK--RGDVIDYS 1405  
 1550 DSDPQOCDSTITSYKGEA--GKFEAN-----GSHTETPEARTKSYFPESONDVYGOS 1601  
 QY 1406 S-----SGVSTNAPSLDPTTEDEKSDQSGSKL-----PGKSSERSSLF 1447  
 1602 TKETLKPKIHSGHVEEPASPLAAY-----OKSLETSTKLITEERKPCVPDLKDSBSSSS 1657  
 QY 1448 QTDLKSGSLRYOKLPSDED-----ESGTEESDNTPLLDKDK-----DRKAEKVERVP 1497  
 1658 EERREVTYRIVIRRLIIGGEAKNIPGESVTEQ-----FTDEGNLITTRKGEQGFVKTK 1712  
 QY 1498 KSPFH 1502

Db 1713 KEIRH 1717  
 RESULT 6  
 US-09-835-788A-12  
 ; Sequence 12, Application US/09835788A  
 ; Patent No. US20020077458A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NI et al.  
 ; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides  
 ; FILE REFERENCE: PT018P1  
 ; CURRENT FILING DATE: 2001-04-17  
 ; PRIOR APPLICATION NUMBER: US/09/835,788A  
 ; PRIOR FILING DATE: 2000-10-17  
 ; PRIOR APPLICATION NUMBER: PCT/US00/28666  
 ; PRIOR FILING DATE: 2000-10-17  
 ; PRIOR APPLICATION NUMBER: 60/159,585  
 ; PRIOR FILING DATE: 1999-10-18  
 ; PRIOR APPLICATION NUMBER: 60/167,246  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 12  
 ; LENGTH: 740  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-835-788A-12

Query Match 4.98; Score 437; DB 10; Length 740;

Best Local Similarity 30.3%; Pred. No. 3,4e-18;

Matches 147; Conservative 66; Mismatches 192; Indels 80; Gaps 13;

QY 6 SOSVINYVEENIPALKALKKCKDVERNECG-QPPLMAAOGNLEIYKEILKNGANC 64  
 303 SQTSNVNASQSMRPYPSV-----DIDARTESNHDPLTLACAGGHEELVSVLIARDAKI 357  
 QY 65 NLEDLNDMTALISAKSGEYHIVEELKCGVNE--HRDMGWTALIMACYKGTDVVELL 123  
 358 EHRDKKGFPLLLAATGAGHYVEILLDKGDIENASERTKOPPLSLACSGGQEVVDLL 417  
 QY 124 LSHGANPSVGLQTSYPTTMAAGRGHADIIVHLLONGAKV--CSDKYGTTPLYAARK 181  
 418 LARGANEHRNVS--DYPLSLAASGGYVNIKILLAGAEINSRTSGKIGISPLMAAMN 476  
 QY 182 GHLECYKHLAMGADV--OEGANSMTALIVAVGGTOSYKELTKRNPVNLTDKQGNRA 240  
 477 GHVPAVKLLDMSDINAQIETNRNTALTACQGAEEVYSLLDKRAVNHRAKTGLTP 536  
 QY 241 LMIASKGEHTEIVODLDACTVYNIP--DRSGDTVLIGAVRGHVEIVALLQKYADIDI 298  
 537 LMEASGTYAEVGRVLLDKGADVNAAPPVPSRDTALTIAADKHYFCELLHGRANHIDV 596  
 QY 299 RGQDNKTALVWAVEKGNATVVRDILQCNPDTEICTGDEGTEPLKATKMRNIEVVELLDK 358  
 597 R-----NKKGVTPMLASNGHFPDVVOLLVOA 623  
 QY 359 GAKYSAVDKGDPTPLHAIIRGSRKLAELLRNPKXGRLLYRNKAGETPVNIDCSHQKS 418  
 624 GADVDAADRNTKTPPLMSAFKGVKYVQYLVK-----EYNQPSDIEC----- 666  
 QY 419 ILTQIFGARHLSPTEIDGMDLGYDLYSSALADILSEPTNQPPICVGLVYQWMSGKSFLLK 478  
 667 -----MRYIA--TIIDKELL--KKHCQCVETIVAKKQDA-----AENAKNMSILLK 709  
 QY 479 KLEDE 483  
 Db 710 ELDLIE 714

RESULT 7

US-09-908-711-70  
; Sequence 70, Application US/09908711



CONDUCT ORIGINATE ADDRESS: /  
/ ADDRESS: Klauber & Jackson  
STREET: 411 Hackensack Avenue 4th Floor



Db 465 NCHGSAVDMAPPELRELYEFGKSLLOAAREADLAKVKKTLAETINEKOPQSHET 524  
QY 134 GLOYSV-----YPIIWAARGHADIVHLLONGAKV 164  
Db 525 ALHCAVASLHPKRYOTELLRRKANVNEKNDFMPLVAAERAHNDVMEVLLHKGAKM 584  
QY 165 NCSDKGTPTVWAARKGHEVCYKHLAMGADVDGANSMTA-----207  
Db 585 MALDITGOTLHRAALAGHLQTCRLLSYSDPSIISLOGFPAQMGNEAVOQILSESTP 644  
QY 208 -----LIYAVKGYQSYKELLRPNVNLMDKOS--NFALMISKEGHTIYODL 256  
Db 645 IRTSDVRYLLASRAGDEIVYKQLOS--SONVNCBDEGRHSTPLHFAAGVYRSVVEL 703  
QY 257 LBAGYVNIPIRSGDVTLLGAVRGHVEIVRALLOKADIDIRGDNKTALYWAVERG-- 314  
Db 704 LHHGADVNAKDKGLVPLHNAACSYGHEVAELLVHGASVNVADLMKTPPLHMAAKKY 763  
QY 315 -----NATWVDIIO-----C 325  
Db 764 EICKLLKHGADPTKKNRDNTPDLVKEGDPDIODLLKGDAALLDAKKGCLARVQKLC 823  
QY 326 NPDEICTKD---GEPPLIKATKMRNIEVEVELLDKGAVSVDKKGDPPLHIAIR----- 378  
Db 824 TPENINCRPTOGNRNTPPLHLAGYNNLEVAETLLEHGADVNAODKGLIPLHNAASYGCC 883  
QY 379 -GRSKRLAELLRNPKDGRLLYRPNKAGETPYNIDC 413  
Db 884 IARVQKLC-----TPENINC 898

RESULT 12  
US-09-972-115A-6  
; Sequence 6, Application US/09972115A  
; Publication No. US20030032769A1  
; GENERAL INFORMATION:  
; APPLICANT: Geron Corporation  
; APPLICANT: Gregg, Morin B.  
; APPLICANT: Walter, Funk D.  
; APPLICANT: Mieczyslaw, Piatyszek A.  
; TITLE OF INVENTION: A Second Mammalian Telomerase  
; FILE REFERENCE: 080/003C  
; CURRENT APPLICATION NUMBER: US/09/972.115A  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: US 60/128,577  
; PRIOR FILING DATE: 2000-04-10  
; PRIOR APPLICATION NUMBER: US 60/129,123  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 1166  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-972-115A-6

Query Match 4.2%; Score 370; DB 9; Length 1166;  
Best Local Similarity 25.6%; Pred. No. 7.4e-14;  
Matches 143; Conservative 61; Mismatches 169; Indels 186; Gaps 14;  
QY 19 PALKALLEKCKD-----VDERNECGO--TPIMIAEOGNLEIVELIKNGA 62  
Db 23 PAARELEFACRNGDVERVKRLVTPPEKVNSRDTAGRKSTPLHFAAGFGRKDYVEYLLONGA 82  
QY 63 NCNLEDLDNWTALISAKSEGHVHVEELKCGVNLHRDMGWTALMAGCYKGRDVEL 122  
Db 83 NVQARDGGGLIPLHNAACSFHAEVYVNLRLRHGADPNARDNMNYPPLHMAIKGKIDVCIV 142  
QY 123 LLSHGAMPV-----TGLOQ 137  
Db 143 LLOHGAEPTIRNDGRTALDLADPSAKAVLTGEYKDELLSARSNGEEMMALITPLNV 202  
QY 138 SVY-----PIIWAARGHADIVHLLONGAKVNCSDK-----YG----- 171

Db 203 NCHASDGRKSTPLHLAGYNNKRYTOVLLDHGADVNAKDKGDLVPLHNAACSYGHEYTEL 262  
QY 172 -----TPVWAARKGHEVCYKHLAMGADVDGANSMTALIVA----- 211  
Db 263 LVKHGACVNAAMDLOFTPLHNAASKNREVCSSLLSYGADPTLLNCHNSAIDLAPPLQ 322  
QY 212 -----VKG--GYTQSYKE-----222  
Db 323 KERATYEFKHSLLQAREADVTRIKHLSLEWNEKHPOTHEALTHCAASPYPRKQI 382  
QY 223 ---IKRNPVNLTDKDGNTALMTASKEGHTIYODLLDAGYVNIPIRSGDVTLLGAVR 279  
Db 383 CELLRGAGNINTEKTEFLPLHNAASEKANDVEYVYKHAENVMLDNIQOTSIRAAV 442  
QY 280 GHVEIYRALLQKADIDIGODNKALYAAVER-----GNATWVDIIOCPN- 327  
Db 443 CGHLQTCRLLSYGDENIISLOGFPAQMGNEVOLLQEGISLGNSEADROLLEBAKA 502  
QY 328 -DTE-----ICT-----KDEG---PPLIKATKMRNIEVEVELLDKGAVSVDKKGDP 372  
Db 503 GDEYTVKLCCTVOSVNCRDIEGROSTPLHFAAGVYRSVVEYLLDHGADVNAKDKGLV 562  
QY 373 LHTAIRGRSKRLAELLRN 391  
Db 563 LHNACSYGHEVAELLVKH 581

RESULT 13  
US-10-163-587A-15  
; Sequence 15, Application US/10163587A  
; Publication No. US20030096263A1  
; GENERAL INFORMATION:  
; APPLICANT: Oliveira, Marcos  
; TITLE OF INVENTION: SELECTIVE PARP-1 TARGETING FOR DESIGNING CHEMO/RADIO SENSITIZI  
; FILE REFERENCE: 50229-306  
; CURRENT APPLICATION NUMBER: US/10/163,587A  
; PRIOR FILING DATE: 2003-01-10  
; PRIOR APPLICATION NUMBER: 60/296,110  
; PRIOR FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 1166  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-163-587A-15

Query Match 4.2%; Score 370; DB 9; Length 1166;  
Best Local Similarity 25.6%; Pred. No. 7.4e-14;  
Matches 143; Conservative 61; Mismatches 169; Indels 186; Gaps 14;  
QY 19 PALKALLEKCKD-----VDERNECGO--TPIMIAEOGNLEIVELIKNGA 62  
Db 23 PAARELEFACRNGDVERVKRLVTPPEKVNSRDTAGRKSTPLHFAAGFGRKDYVEYLLONGA 82  
QY 63 NCNLEDLDNWTALISAKSEGHVHVEELKCGVNLHRDMGWTALMAGCYKGRDVEL 122  
Db 83 NVQARDGGGLIPLHNAACSFHAEVYVNLRLRHGADPNARDNMNYPPLHMAIKGKIDVCIV 142  
QY 123 LLSHGAMPV-----TGLOQ 137  
Db 143 LLOHGAEPTIRNDGRTALDLADPSAKAVLTGEYKDELLSARSNGEEMMALITPLNV 202  
QY 138 SVY-----PIIWAARGHADIVHLLONGAKVNCSDK-----YG----- 171  
Db 203 NCHASDGRKSTPLHLAGYNNKRYTOVLLDHGADVNAKDKGDLVPLHNAACSYGHEYTEL 262  
QY 172 -----TPVWAARKGHEVCYKHLAMGADVDGANSMTALIVA----- 211  
Db 263 LVKHGACVNAAMDLOFTPLHNAASKNREVCSSLLSYGADPTLLNCHNSAIDLAPPLQ 322  
QY 212 -----VKG--GYTQSYKE-----222

Db 323 KERLAEEKSHSLLOARREADVTRIKKHLSLEWVNEKHPOTHEHTALHCAASPYPRKQI 382  
 QY 223 ---ILKNPVNLTDKDGNALMIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVR 279  
 Db 383 CELLERKANINETKEFLPLHVASSEKANDVEYVVKHAEVNMALDNGQSLHRAAY 442  
 QY 280 GGHVEIYRALLQYADIDIRGODNKALYNAVEK-----GNATVRILOCPN- 327  
 Db 443 CGHLOTCRLLSYGCDENIISLOGFTALQNGNEVVOOLOEGISLGNSEADROLLEAKA 502  
 QY 328 -DPE-----ICD-----KDE-----PLIKATKMRNIEVELLDKGAKEYSAVDKGDTP 372  
 Db 503 GDEYVYKLLCTVOSVNCRDIEGROSTPLHRAAGINRSVVEYLLQHGADVHADKGSIVP 562  
 QY 373 LHAIRGRSRKLAELLRN 391  
 Db 563 LHMNCSYGHVEVALLVKH 581

RESULT 14  
 US-09-509-196A-2  
 ; Sequence 2, Application US/09509196A  
 ; Patent No. US20020037582A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DAIX, Roger J.  
 ; APPLICANT: SUTHERLAND, Robert L.  
 ; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling  
 ; FILE REFERENCE: 1871-129  
 ; CURRENT APPLICATION NUMBER: US/09/509,196A  
 ; PRIOR FILING DATE: 1997-09-23  
 ; PRIOR APPLICATION NUMBER: P09388  
 ; PRIOR FILING DATE: 1997-09-23  
 ; PRIOR APPLICATION NUMBER: P09388  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 1074  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-509-196A-2

Query Match 4.0%; Score 357; DB 10; Length 1074;  
 Best Local Similarity 23.9%; Pred. No. 4e-13;  
 Matches 150; Conservative 77; Mismatches 211; Indels 190; Gaps 13;  
 QY 14 EENIPALKALLECKDVERNEGOTPLMTAEOGNLEIYKELIKNGANCNLEDDDMWT 73  
 Db 98 EKKMALLTPLNVNCHADGRK---STPLHLAAGYNRYKIYOLLQHRDVAHDKGDLV 154  
 QY 74 ALISAKEGHVIYELLKCGVNIENRDMGNTALMAACYGRDVLVELLSHGANSPT 133  
 Db 155 PLHNACSYGHEVEYELLYKHGGCVNAMDLMOFTPLHEAASKNRVEVCSLLSYGADPTLL 214  
 QY 134 G-----LOYS----- 138  
 Db 215 NCKNKSALDAPTPOLKERLAEFEKGSLLQARADVTRIKKHLSLEWVNEKHPQIHT 274  
 QY 139 -----YPLITMAAGRHADIVHLLONGAKV 164  
 Db 275 ALHCAASPYPRKQICELLERKANINETKEFLPLHVASSEKANDVEYVVKHAEV 334  
 QY 165 NCSDRYGTPLVMAARKGHEVCVHLLAMGADV-----QEG-- 201  
 Db 335 NALDNLGOTSLHRAAYCGHLOTCRLLSYGCDPNIISLOGFTALQNGNEVVOOLOEGIS 394  
 QY 202 ---ANSMPTLIYAVVGCTOSYKE----- 224  
 Db 395 LENSEADROLLEAKAAGVEYVKKCTVOSVNCRDIEGROSTPLHRAAGINRSVVEYLL 454  
 QY 225 KKNPVNLTDKDGNALMIASKEGHEIYODLLDAGTYVNIIPDRSGDTVLIGAVNGHVE 284

Db 455 QHGADVHADKGSIVPPLHNAACSYGHVEVALLVKHGAAYVNNADLKKFTPLHEAASKYE 514  
 QY 285 IYRALLQYADIDIRGODNKALYNAVEKGNATVRILO----- 324  
 Db 515 ICKLLQHGADVPPKKNRDENTPL--DLVKGDTIDIDLLRGDAALLDAKKGCLARVKL 572  
 QY 325 CNPDPEICCKDE---PLIKATKMRNIEVELLDKGAKEYSAVDKGDTPPLHRAIRGR 380  
 Db 573 SSPDNVNC-RDQGRHSTPLHLAAGYNLEVAEYLLQHGADVHADKGSIVPPLHNAASYG 631  
 QY 381 SRKLAELLRNPKRDEHLYRPNKAGETPYNIDCSHOKSILITQIFGARHLSPT-----ET 434  
 Db 632 HDVVAALLI---KYNASLWATDKMAFTPLHEAAGKRGTOQLALLAHGADPPLKMQEGOT 688  
 QY 435 DGDMLGYLYSSALADILSEPTMQPPTC 462  
 Db 689 PLDLVSAVDVSAALLTAA--PPSALPSC 714

RESULT 15  
 US-10-164-080-7  
 ; Sequence 7, Application US/10164080  
 ; Publication No. US2003008741A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BIRD, Timothy, A.  
 ; APPLICANT: HOLLAND, Pamela, M.  
 ; APPLICANT: PESCHON, Jacques, J.  
 ; APPLICANT: VIRCA, George, D.  
 ; TITLE OF INVENTION: DEATH ASSOCIATED KINASE CONTAINING ANKYRIN REPEATS (DAKAR) AND  
 ; FILE REFERENCE: 3280-B  
 ; CURRENT APPLICATION NUMBER: US/10/164,080  
 ; PRIOR FILING DATE: 2002-06-04  
 ; PRIOR APPLICATION NUMBER: 60/295,959  
 ; PRIOR FILING DATE: 2001-06-04  
 ; PRIOR APPLICATION NUMBER: 60/334,362  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 7  
 ; LENGTH: 784  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-164-080-7

Query Match 3.98%; Score 348.5; DB 9; Length 784;  
 Best Local Similarity 27.7%; Pred. No. 8.3e-13;  
 Matches 99; Conservative 64; Mismatches 126; Indels 69; Gaps 5;  
 QY 35 NEGOTPLMTAEOGNLEIYKELIKNGANCNLEDDDMWTALISAKEGHVIYELLKCG 94  
 Db 468 NRGSTPLHMAVERVAGVYELLARKISYNANDEOQMTLHRAQNGDSSSTRLLLEN 527  
 QY 95 VNLHRDMGNTALMAACYGRDVLVELLSHGANSPTGLQYVYPLITMAAGRHADIV 154  
 Db 528 ASVNEVDGEGRTPHNVACQHGQENIVIRILLRGVDSIQ--KDAWPLRPLHAAQGHILPY 586  
 QY 155 HLL-ONGAKVNCSDKGTTPVMAARKGHEVCVHLLAMGADVDOEGANSMTALVAYK 213  
 Db 587 KILAKQGVSYNAOTLDGRPLHLAAGRHVARILIDLCSDVNNCSLLAOTPLHNAE 646  
 QY 214 GGYOSVKEILKRPNVNLTDKDGNTALMIASKEGHEIYODLLDAGTYVNIIPDRSGDV 273  
 Db 647 TGHSTARLLHRAAGKAMTSOQYTAHLAA----- 678  
 QY 274 LIGAVRGHVEIYRALLQYADIDIRGODNKALYNAVEKGNATVRILOCPDPEICT 333  
 Db 679 ---RNGHLATVLLVEERADVILARPLNOTALHLAAAGHSEVELEV----- 723  
 QY 334 KDGETPLIKATKMRNIEVELLDKGAKEYSAVDKGDTPPLHRAIRGRSRKLAELLRN 391  
 Db 724 -----SADVIDLE-----DEQGLSALHLAAGHQAQVETLLRH 757



Tue Jul 1 15:13:30 2003

us-10-021-571-4.rapb

Page 10

Search completed: July 1, 2003, 14:51:10  
Job time : 51 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:05 ; Search time 19.5 Seconds

(Without alignments)  
2587.708 Million cell updates/sec

Title: US-10-021-571-4

Perfect score: 8884

Sequence: 1 MSYLIQSQSYNTVEENIPA.....LHMAASSESTGGEHRENTL 1715

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfillseq1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	530.5	6.0	1839	2	US-09-172-977-4
2	493.5	5.6	1843	2	US-09-172-977-3
3	480.5	5.4	1745	2	US-09-031-485-33
4	480.5	5.4	1745	2	US-08-847-429A-33
5	480.5	5.4	1745	3	US-09-065-474-33
6	480.5	5.4	1745	4	US-09-557-034-33
7	438	4.9	1088	4	US-09-082-059-2
8	409.5	4.6	1423	4	US-08-810-712-10
9	376.5	4.2	1327	4	US-09-196-387-2
10	374	4.2	348	2	US-09-031-485-28
11	374	4.2	348	2	US-08-847-429A-28
12	374	4.2	348	3	US-09-065-474-28
13	374	4.2	348	4	US-09-557-034-28
14	370.5	4.1	949	4	US-09-196-387-10
15	363	4.2	1166	4	US-09-350-982C-5
16	356.5	4.0	352	3	US-09-065-474-139
17	356.5	4.0	352	4	US-09-557-034-139
18	353	4.0	302	2	US-09-031-485-38
19	353	4.0	302	2	US-08-847-429A-38
20	353	4.0	302	3	US-09-065-474-38
21	353	4.0	302	4	US-09-557-034-38
22	353	4.0	303	2	US-09-031-485-23
23	353	4.0	303	2	US-08-847-429A-23
24	353	4.0	303	3	US-09-065-474-23
25	353	4.0	303	4	US-09-557-034-23
26	353	4.0	741	2	US-08-436-771-4
27	353	4.0	741	2	US-08-434-998-4

28	353	4.0	741	2	US-08-487-797-4	Sequence 4, App11
29	353	4.0	741	5	PCT-US95-02058-4	Sequence 4, App11
30	348	3.9	679	2	US-08-462-481-4	Sequence 4, App11
31	348	3.9	679	2	US-08-436-771-6	Sequence 6, App11
32	348	3.9	679	2	US-08-434-998-6	Sequence 6, App11
33	348	3.9	679	2	US-08-487-797-6	Sequence 6, App11
34	348	3.9	679	2	US-08-701-005A-4	Sequence 4, App11
35	348	3.9	679	2	US-08-479-895-4	Sequence 4, App11
36	348	3.9	679	5	PCT-US95-02058-6	Sequence 6, App11
37	344	3.9	679	3	US-08-943-956A-4	Sequence 4, App11
38	344	3.9	787	4	US-09-188-930-334	Sequence 334, App
39	343	3.9	741	2	US-08-462-481-2	Sequence 2, App11
40	343	3.9	741	2	US-08-436-771-2	Sequence 2, App11
41	343	3.9	741	2	US-08-434-998-2	Sequence 2, App11
42	343	3.9	741	2	US-08-487-797-2	Sequence 2, App11
43	343	3.9	741	2	US-08-701-005A-2	Sequence 2, App11
44	343	3.9	741	2	US-08-479-895-2	Sequence 2, App11
45	343	3.9	741	5	PCT-US95-02058-2	Sequence 2, App11

## ALIGNMENTS

RESULT 1									
US-09-172-977-4									
; Sequence 4, Application US/09172977									
; Patent No. 5989863									
; GENERAL INFORMATION:									
; APPLICANT: Tang, Y. Tom									
; APPLICANT: Guegler, Karl J.									
; APPLICANT: Corley, Neil C.									
; APPLICANT: Yue, Henry									
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN									
; FILE REFERENCE: FF-0615 US									
; CURRENT APPLICATION NUMBER: US/09/172, 977									
; CURRENT FILING DATE: 1998-10-14									
; NUMBER OF SEQ ID NOS: 4									
; SOFTWARE: PERL Program									
; SEQ ID NO 4									
; LENGTH: 1839									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
; FEATURE:									
; OTHER INFORMATION: 929491									
US-09-172-977-4									
Query Match									
Best Local Similarity 20.2%; Score 530.5; DB 2; Length 1839;									
Matches 426; Conservative 263; Mismatches 654; Indels 763; Gaps 81;									
QY	14	EEENIPALKALKECKDVEDERNEGOTPTMAAEGCNLEIVKELIRNGANCLEIDDMNT	73						
DB	73	KEEHVGLVOLLRLRGSSVDSATKGTALHIASLAQAEVYKLVVEGANINAGSNGFT	132						
QY	74	ALISAKKEGHVHIVEELKCGVN-----LEHRDYG---	103						
DB	133	PIYMAAGENHIDVYKLENGANOSTATEDFTPLVALQGHNAQVAILEEDTQKVR	192						
QY	104	-----GWTALMWACYKGTVDVVELL	123						
DB	193	LPALHTAARKDRTKSAALLLONDHNDVOSKMNVRFTSGFTPLIAAHYGVNVAATL	252						
QY	124	LSHGANSVYGLQSYVPTIIMAGRGHADIVHLLONGAKVNCSDRYGTTPLVMAARKS-	182						
DB	253	LNRGAADVFTA-RNGITPLHVAASKRGTNTNKKLLDRGQIDAKTRDGILPLHCAARSH	311						
QY	183	-----HLECYKHLANGADVDGANSMTALV	210						
DB	312	DQVVELLEGGAPLARTKNGLSPLHMAAGDHEVCVKHLIDQKAPVDVTLDTLTAHV	371						
QY	211	AVKGGYQSKEIKLRPNVNLTKDQGNLTMAKSE-----	247						
DB	372	AAGGHYRVTKLLDKRANPNARALNGFTPLHACKKNRIKIKWELLVYKGASIQAITESG	431						

QY 248 -----GTEIVODLLDAGTVNIPDRSGDVLIGAVGHEVEIRALLQKADID 297  
 DB 432 LEPHYAAMGLNLTIVLLLOMGASPDVTNRIGETALMMARAGVEYVRLIRGALVLD 491  
 QY 298 IGCQNKALYAVKGNATVNDILQCPTEICTKDEPPLIKATMRNIEVEVELLD 357  
 DB 492 AARBEQTPHIASTRIGTEIYVOLLQMHAPDADATNGYTPHISAREGQVDAVSLLE 551  
 QY 358 KCAKSAVDDKGDPTLHAIGRSRKLAELLR-----NPKD 394  
 DB 552 AGAASLTKKGFTHLAAKYGSLDVAKLLOPRAADSAGKGLPLHVAHYDMQKV 611  
 QY 395 GRL-----YRPNKAGETPYNIDC-SHOKSILTOI--FGAR-----HLSP 432  
 DB 612 ALLLEKASPAHATKNGYTPHIAKKNQMOGIASTLLNGAEFNIVTKOGVTPHLSAQ 671  
 QY 433 EDDGML-----GVDLSSALADLSEPTMOPPCVLYAQMGSGSKFLKLKLEDEK- 484  
 DB 672 EGHDMVTLLDKGNIMHSTKSGLTS-----LHLAAO-----EDKYNV 710  
 QY 485 -----KTPAGOOIEPLFOFWMLIVELTLILC--GGLGL--PFTVHPMLIGIAVSLFLAL 536  
 DB 711 ADILKKGADDADHAKLGLTPLY-----ACHGVKKNVNFLLQGANVNAKTKNGYTP 765  
 QY 537 -----LYFYIYIFGGRREGESNMWAVLSTRARHIGYLEL--IKLM----- 579  
 DB 766 HQAAQOQHTHINVLQHGAKPNATYANGNTALA--IAKRLGYISVDTLKVTEEVYTT 823  
 QY 580 -----FVNPELPEOTKALP--VRPL-F 600  
 DB 824 TTTITEKKNVPEMTVEVLDVSEDEDDTMTGGGGEYLRPEDELDKEDDLSPPSOPLDG 883  
 QY 601 TDYNRLSSVGET--SL-----AEMTILSDACERE 629  
 DB 884 MNYLRYSLDEGRSDLSRFSFSDRSHTLSHASYLSDAAMDVSVIPSHQVSTLAKEMRN 943  
 QY 630 EGFALTRFVRFKIEDVOGKKKKKKTCLBSFYFLP-----IIGCISITILLAFRV 683  
 DB 944 SYRLS-----WGTEINDVALSSSPIHSGFLVIFMVARGARCGRHNGLRII--I 993  
 QY 684 DPKHLTVNAVLIASVYGLAFVLCRTMQLVDSLSLNSOKRL-----HNAASKL 734  
 DB 994 PPRCTAPT-----RYTCRL-----VKHRLATMPRVEGGLASRL 1030  
 QY 735 HKLSEG-----FMVKLCEVELMAMARTIDSTFON-----QTRLVVIIDG 776  
 DB 1031 IEVPSGAGFLGPIYVEIPHPAALRGKERELVLRSENGDSMKHEHPCDYDEDELINELNG 1090  
 QY 777 LDAGEODKVLQMDTVAVLPSKGFIAIFASDPH--IIRAINQNLNSVLRDSINGHDY 834  
 DB 1091 MD-----EVLDSPEDELEKRIKRIITNDPQYFAVNSRIKO-----DSNLIGPE- 1134  
 QY 835 MRNIVHLPVFLNSGLSNARKFLVTSATNGDVPSCDITGQIODEADRRVSONSLGEMTK-- 892  
 DB 1135 -----GGVLSSTVVPQOVAVPEGALTKRI--RVGLQ-----APMSELYKTI 1176  
 QY 893 LGSKTALNRDITY--RRQOMQRTITROMSFDLTKLVTESWSDISFQMRRLINIVSVT 950  
 DB 1177 LGNKATESPIYTLTPRRKRKHKPIITMTIPVKASSDVMWLGFGDAP--TIRLL--CSIT 1232  
 QY 951 GRLRANOISFNMRLASWMLT--EQWPYRTS-----WLIILETBEBIPOMLTKI 1002  
 DB 1233 GGTTPA-----QWEDITGTPPLTFVNECVSTTNVSAKFWLI--DCROIQESVTFASQ 1283  
 QY 1003 YERISKNIPTTKDVEPLEIDGDIRNEFVLSRTPLVARDVKVLPCTVNLDPKL-- 1059  
 DB 1284 VYREITGVY-----MAKVVFPAKSHD-----IARLRRCMCDKVDYKT 1324  
 QY 1060 ---REITADYARAREQISIGLAVPLPLHEGPPRADSGVSOQPSVCSSTFNGPPAGV 1116  
 DB 1325 LEODENFAVARSHD-----VEVLEGRPIYVDCF-----GNL 1356

QY 1117 V-----SPQHSYSSGWTGPQHPFYNRGSGPAPGVVILNSLVNDAVCERKIQIEGDS 1172  
 DB 1357 VPLTKSGCHHIFSEFAKREKRLPLFVKVRDTPQPCGRUSFM-----KEPSTRGLVHQ 1410  
 QY 1173 MLPOYCTTIKANINGRVLAQCNID-----ELKEMNMNFGDHLFRSTVL 1218  
 DB 1411 -----AICNLNTLPIYTKESSEDDQOEEDIMT----- 1439  
 QY 1219 EMNAESHVVPEDPRELSESSGPPADHGPAPARASINELPHLELSQTPYTLNFSFEELN 1278  
 DB 1440 -----SEKNPODEORIEB-----RLATYADILGFSWELARE--LDFTEEQJH 1481  
 QY 1279 TLGDEGAPRSHNLSMQSO-----TRRPPSSLSNDSODSIEISKLPDK 1322  
 DB 1482 QIIE-----NPNLOQSOQLKIMLEERQKHAITDNLVCECLKIRMDIYHLMETNERP 1537  
 QY 1323 VQAEYRDAYREYIAQMSQLEGGFSTTIGR--SSPHSTYWGQSSSGSIHNSLEQEK 1380  
 DB 1538 LOERISHSYAE--IEQITLIDHSEGFSLQOEELCTAQHK-----QKEBOAVS 1582  
 QY 1381 KDEPFPDQGRKSFLLKRGVDIDSSSGVSTNDASPLDITPED-----EK 1426  
 DB 1583 KESE-----TCDHPI--VSEBDSVGTSTQDGVPKTE 1614  
 QY 1427 SDQSGSKLLPGKKSERSLFTDLKIGSGRYOKLPSPDED--ESGTERSDMTPLKDK 1485  
 DB 1615 GDSSSTALFPQTHKEQ--VOODF--SG-KMODLPRESSLEFQOEFTVTPGTETSE 1665  
 QY 1486 DRKAEGKVERVPSPEHSAPIRTFIKAKEYLSDALDKKSDSDGVRSESSPMHSLN 1545  
 DB 1666 TOKA--MIVSPSPKTE----- 1681  
 QY 1546 EVADDSQLEKANIIELEDSDHSGKRGIPHSLSGLQDPIIARMSICSE--DKKSPSCSLI 1603  
 DB 1682 EVSTPADEEK--LYIQTPSSERG--SPIQEPPESEHREESSPKRTSLV 1729  
 QY 1604 ASSPEENWP-----ACQKAYNLNR--TSTYTLNNSAPARANONFDEMGEIR- 1651  
 DB 1730 IYESADNOPETCERLEDDAFAFEKDDMPEIPEYV-----EEEDIDEH 1773  
 QY 1652 -----TSOYLRRSS-----PNTTIONMLSMTHKRSQBSYRL 1689  
 DB 1774 GHTVVKVTKRIIRRYVSEGEKEEIMVQGMPOEIVNEEDGYSKYIKRVYLSKSDTEQ 1833  
 QY 1690 SKDPPE 1695  
 DB 1834 SEDNNE 1839

RESULT 2  
 US-09-172-977-3  
 ; Sequence 3, Application US/09172977  
 ; Patent No. 5969863  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Guegler, Karl J.  
 ; APPLICANT: Corley, Neil C.  
 ; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
 ; FILE REFERENCE: PF-0615 US  
 ; CURRENT APPLICATION NUMBER: US/09/172,977  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 3  
 ; LENGTH: 843  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: g1841966  
 US-09-172-977-3

Query Match 5.6%; Score 493.5; DB 2; Length 843;



```

QY 392 -----PKDRLLRPNKAGETPYNIDCSHOKSILNIOIGARHLSPETIDGMLG 440
DB 686 QEDRVSAEELVKENAAIDPKTKAGYTPLVHACH-----IG-----QINMVR 727
QY 441 YDLSSALADILSEPTMPPICVGLIQAOWSGSKSLKLEDEKMTAGQOIEPLFQPSW 500
DB 728 FLIEHGANVSITRASYTP-----LHOAAOQHNSVVR-----760
QY 501 LIVETLLLCGILGLFAFVHPNGLIANSLSFLALTYIFLIYVFGRRGESMNAMV 560
DB 761 -----VLEHGASPNVHTSGOTPLS-----781
QY 561 LSTRARIHIGLELLIKLMYVNPBLPOTTKALPVRE-----LETDYNRLSV 609
DB 782 -----IAERIGYVSVEALKITETTVITETTVTEERKPCPNPAMETMMSD-----SED 833
QY 610 GGETSLAMIMTILSPACEREGFL--ATRLFRVKTEDTQCK-----650
DB 834 EGEDNQTAMAHADFESLTKGLHSDGVHLHATEPTLSRSEVEGTDDLDALIRKA 893
QY 651 -----KWKTCCLPSFYIFLEI-----IGCIISGITLLAIF 681
DB 894 QHEPITAMADPSIDASLPDNTVIMRTMQPSFLISPMVADARGAMGCRHSYRIIIP 953
QY 682 RVDRKHLTVANVLISVGLAVLNCRTMWOVLDSILN-----SQRRLHNAASKLHL 737
DB 954 RKAQOPTRV-----TCR--YLCRKLAPHPPLSEGEAL--ASRILEM 991
QY 738 KSEG--FMKVLKCEVELMARMAKTIDSETONOTRLVYII--DG-----LQACEDKV 765
DB 992 APHAKELGVYILEVPHASL-----KREKEIYILNSDDQHKEHOLEVTE--DAV 1042
QY 786 LQMLD-----TVRLFSKGFPIAFASDPHIIKAINQNLNSVLRDS 827
DB 1043 QEVINESFDAEELSOLDLHTRITRLITNDPFI--YEA-----VYTRVQEVHCVQPEG 1095
QY 828 NINCHDMRNIV--HLPEFLNSRGLSNARKFLVTSATGDVPCSDTGTIOEDADRVSONS 886
DB 1096 GV-----ILSSVPHVQALPPDGSIT--KIKKVSQAQPPQOELVYRLH-----1137
QY 887 LGEMTKLGSKTALNRDVTY--RRROMQRTTRROMSFDLTK--LVTYEDWFSISPQTM 941
DB 1138 -----GNRVAASPIYVEPRRRKFKPITLICILPQSNKMLVOYSQPOQEPPTLR 1190
QY 942 RLL-----NIVSVTGLRLRANOISEN-----W-----DLASWI--N 971
DB 1191 LLSKGTGSSPAQWEDITGTOTLFTGEDVSTTVSARFWMDCOTPRDARMAQEVYN 1250
QY 972 LTEQWPTYSMLILYEETEGIPDOMTL-----KIT--YERISKNIPTTKDVEPL- 1019
DB 1251 EALAVPYMAKFLI-FARTRPAGQLRLFCMTDREDKYLEKORFTE--TAKSKDEVILS 1308
QY 1020 -----LEIDGDIRNEFEVLSRRPVAVARD--VKVPLPCTVN--LDPKLR-----EIIA 1064
DB 1309 GRHOFEFFSGNL-----LPITKSGDQLSLYFLPQENRLAMVIRHTDNETAA 1358
QY 1065 DVBAARQOISIGGLAVPLPLHGGPRRABGYSQPVSQSSSTFNPFRFAGVGVSPOPHS 1124
DB 1359 DGR-----IYMKERKLAENLPPQTP-----VCTLATITLPEYTG--PEPMVS 1399
QY 1125 ---YYS--GNTGPOHFFYNGSGPAPGP--VVLNLSLVNDVACEKIL-----Q 1165
DB 1400 KKLFESEASLLEKYGAFHETAPROMLPLAHVALLLGAOMHRLARALEVPDIDIROVHRQ 1459
QY 1166 IEGLDQSMPLROYTTTIRKANIN--GRLVLAOCNIDELKEMNMN-----F 1207
DB 1460 LVLELVAVTILRLIWLFLKKEQAPVALRSALORIGRDVVRREMDRAKLDGLBTPVSHS 1519
QY 1208 GDHHLFRSTVYLEM-----RNAE-----SHYVPEDPRPLSESSSG--PAPHEEPARRASHNL 1257
DB 1520 GRSITISLLEAVAGRRHHAETMAOQRLAQEPFQOVGYNGTGEDPEPEQSFHEHE 1579
QY 1258 PHTLESSQTPYTLNFSFEELNLTGLDEGAPRHSNLSWOSQOTRRTPLSLNSQDSIETS 1317

```

```

DB 1580 EBAVVS-----EIRIVRTE--RHVHSENGPIYEERTTTTTYEDDVAVNEE 1624
QY 1318 KLTQKV-----QAEYRAYREYIAQMSQEGSPSTTISGRSSPHSYMGQSSGGS 1370
DB 1625 EIVDKYIPLNEEBOEKMDRVAREVENFQOE-----TSKEGFGCOTTHEKKNDDGS 1679
QY 1371 IHSNLEOKKQKDEPRP--DGRKSFLLKRGVDYIDSSGVSTNDASP-LDPTIED 1424
DB 1680 LKTTM-----KSHVHQIFEDGGETS-----ANETGLSSGDADITMPTTKED 1722

RESULT 4
US-08-847-429A-33
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-847-429A-33

Query Match 5.4%; Score 480.5; DB 2; Length 1745;
Best Local Similarity 20.3%; Pred. No. 5.6e-32;
Matches 353; Conservative 234; Mismatches 602; Indels 549; Gaps 67;

QY 14 EENRPLAKALLECKVDNERNEGQTPPLMAAQNLLEYKELIKGANCNLEDLN-- 71
DB 207 KKDPTKATATLLQNDHNSDVTSKSGFTPLHIAHYGNENVAQQLLEKGANVNVQARHNIS 266
QY 72 -----W-----TALISAKRGHVIHYELLKCG-----94
DB 267 PLHVAITKGRNTMYSLLAHAGAVIDCTROLITPLRHCASSGSDQVYDILLLEGAPISAK 326
QY 95 -----VNLH-----RDMGWTALMW 110
DB 327 TKNGIAPLHMAAQVDVTVYITPLHVAACGHVRAKLLLDNRADPNARALNGFTPLHI 386
QY 111 ACYGRGDVYELLSHGANSVYGLQYSYVPIIAGRGADIVHLLDNGAVNCSDKY 170

```

```

Db 387 ACKNRKIKYELLKTHAHEAT-TESG:SLPHVAFMGAINIYIYLLQGANADAVTR 445
QY 171 GTPLVMAKGLLECYKLLHANGADYDDBGANSMTALVAVKGTOSYKELKRNPV 230
Db 446 GETPLHAAANOTDIRVLYRNGAOVDAARELQTPHIAHSLGNTDVIYLLQANASP 505
QY 231 NLTKGNALMAKSGHTEIYQDILDDGTYNINIDRSGDPTLIGAVGHEIYRALL 290
Db 506 NAATRODLYPLHIAEGOEVAAILHMDGTDLTKKGFPLHIAKGNLPAKSL 565
QY 291 OKYADIDIRGODKKTALYMAVEKGNATWVDILQCPDTEICTKDETPLIKATKRNIE 350
Db 566 ERGTVPDIEKNGVPLHVAHAHNDKVALLLENGASAAKKNVTPPLHIAKKNQMD 625
QY 351 VELLIDKAKVAVDKGDTPLHIAHRSKRLAELLRN----- 391
Db 626 IASTLLHYKANANAESAGFTPLHAAOEGHREMAALLENGAKVGAQANGLTPHILCA 685
QY 392 -----PKDRLLYRNKAGFTPNIDCSHOKSILTOJFGARHLSPTEDDMLG 440
Db 686 QEDRVSAEELVKNENAIIDPKTAGYTPHVAACH-----FG-----QINMYR 727
QY 441 YDIYSSALADILSEPTMQPICVGLXAWGSGKSLKLEDEKMTFAGQIPLFOFSW 500
Db 728 FLIEHARVASVITRASVTP-----LHOAQOGHNSVVR----- 760
QY 501 LIYFLITLLCGGLGLFAFTVHRLGIAVSLFALLIYFIYIYGGREGESWMAWY 560
Db 761 -----YLLEHGASPNVHTSTGOTPLS----- 781
QY 561 LSTRARHIGYLELLKLMFVNPELPEDOTKALPVAF-----LFTDYNRLSV 609
Db 782 -----IARLGVSVAELAKTITETVTEETTYEENYKQNDPEAMETMFS-----SED 833
QY 610 GGETSLAEMIALTSDACEREGFL--ATRLFRVFKTIDTQCK----- 650
Db 834 EGEDNOITANAHADFESESLTKGLHDSGTGVLHATEPTLSRSEPEVBDGDDALIRKA 893
QY 651 -----KWKTCCLPSPVIFLFI-----IGCIIISGTLIAIF 681
Db 894 OHEPTTAMADPSLDASLPDNTVMTKRTMOPSLISMVADRGAMGCHRSGRITIPP 953
QY 682 RYDPRHLTVNAVLISISVGLAFVLCRTWMOVLDSILN-----SQRRLHNAASKLHL 737
Db 954 RKAQOPTRV-----TCR--YLCKDKLAHPPLSEBELL--ASRLTEM 991
QY 738 KSEG--FMKVKCEVELMARMAKTIDSTFONQRLVYI--DG-----LDACEODKY 785
Db 992 APHGAKFLGPVILEVPHFASL-----RGEREIYILRSDDGOHMKHEHLEATE-DAV 1042
QY 786 LQMLD-----TVRYLFSKGPFIAPASDPHIIKAINONLSVLRDS 827
Db 1043 QEVINESFDELSQDLDTSLRITRLITNDPEM--YFA-----VTVRROEVHCVPEG 1095
QY 828 NINCHDYARNIV-HLPVFLNSRGLSNARKFLVTSATNGVPCSDTGTIOEDADRVSNS 886
Db 1096 GV-----LLSYVPHVQAIFFDGSLL--KTIKVSVQAPVQELVTRLH----- 1137
QY 887 LGEMTKLGSKTALNRDITY--RRQOMOTTIRQMSFDLTK--LLVYEDWFSDISPOTMR 941
Db 1138 -----GNRVAVSPIVTVEPRRRKFRHPIRLCIPLOSNNKMLTOYSGOPGOEPPLR 1190
QY 942 RLL-----NIVSVTGRLLRANQISFN-----W-----DRLASW-N 971
Db 1191 LLCSKTGSSPAQWEDITGTTOLFTGDEVSFTTVASRFLMCCQTPRDAARAQOEVYN 1250
QY 972 LTEQWPTSWLILYEETEGIPDOMTL-----KTI--YERISKNIPTKDOVEPL- 1019
Db 1251 EAIANVPYAKFLI-FARTPFAEQGLRFLCMTDREKTLKORERFIE-IAKSDVEYLS 1308
QY 1020 -----LEIDGDIRNEVEVLSRTPVLVARD--VKVFLPCYV--LDPKLR-----ELIA 1064
Db 1309 GHOHFLFFSGNL-----LPTKSGDOLSLTFLPQENRLAFVAKIRTHDNETAA 1358

```

```

QY 1065 DVRAAREQISIGGLAYPLPLHEGPPRAPSGVSGPVSQSSSTSENGFAGVSPORHSS 1124
Db 1359 DGR-----IYFMKEPKLRAENLPQTP-----VCTIAITLPEYTG-----PEMVVS 1399
QY 1125 ---YYS--GMTGPQHPFYNRSGPAPG---VVLNSLNVAVCEKLR-----Q 1165
Db 1400 KKLFSSEASLTERKYAGAFHETAEPNLPLAHVALLIGADHRLARALEVPDIDIRQVHQ 1459
QY 1166 IEGLDQSMPLQYCTTIKKANIN---GVLACNIDELKEMNN-----F 1207
Db 1460 LVGLAVITLRIWIPLKREQATPVALRSALORIGDGVREMDRAEKLDLEGTPVSHIS 1519
QY 1208 GDMHLFRSTVLEM-----RNAE-----SHVPEDPRLSESSSG-PAPHPARARASHNEL 1257
Db 1520 GGSITSLSTLLEAGDRRHAEVTAQAQRIADPEFTQOVGNGTGDPEDEKEOSFHEE 1579
QY 1258 PHTLSSQTPYTLNFSFEELNTLGDGAPRHSNLSWOSQTRRPPLSLSSLSQSSSIS 1317
Db 1580 EEVAVS-----EIRTVRTE--RAVHDSNGPIVEERTITTYEDDVAVNEE 1624
QY 1318 KLTDKV-----QAEIRAVREYIAQMSQLEGGGSTTISGRSSPHSTYTMGSSSGS 1370
Db 1625 EIVDKIVPLNEEOEKMDRVRVEVEMNEEQOE-----TSKEGTFCQTTHEKEKDDGGS 1679
QY 1371 INSNLEOEKGDSEKPP--DDGRKSFLMKRGDVIDYSSGSVSTNDASP-IDPTIIEED 1424
Db 1680 LKTTM-----KSHVNAQIFPDGERS-----ANETGLSSDADPTIMPTTKED 1722

```

```

RESULT 5
US-09-065-474-33
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESS: Heskia Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-Apr-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-065-474-33

```

Query Match 5.4%; Score 480.5; DB 3; Length 1745;  
 Best local Similarity 20.3%; Pred. No. 5 6e-32;  
 Matches 353; Conservative 244; Mismatches 602; Indels 549; Gaps 67;

14 EEBNIPALKALLKCKDNERNECGOTPLMTAAEOGNETIVKELIKGANCNLEDDLN-- 71  
 207 KKDPTKAAATLLQNEHNSDVTSKSGFTPLHIAHGNENVAQLLEKGANVYQARHNIS 266  
 72 -----W-----TALISAKGCHVIVELLKCG----- 94  
 267 PLHVATKMGRTMWSLLAHGAVIDCKRTDLTLPLHCASRSGHDQVDDLLEKGPISAK 326  
 95 -----VLEN-----RDMGWTALMW 110  
 327 TNGKGLAPLHMAQVDVVDYDYLPLHVAAGHGVAVKALLDRNADPNARALNGFTPLHI 386  
 111 ACYKGRDVTVELLSHGPNPSTVGIQYVPIYMAAGGHADYVHLLQNGAKYNSCKY 170  
 387 ACKKRIKIVELLTYHAIEAT-TEGSLPLHVAFMGAINIYIYLLQOGAMADVAVR 445  
 171 GTPLVMAARKGHECVKHLAMGADVDEGANSMTALIVAVKGGYTSKEILKRNPNV 230  
 446 GETPLHIAARAQDITVAVLVANGQVDAARELOTPHIASRGNTDYLILLQANASP 505  
 231 NLTDDGNTALMASKEGHEIYDLDAGTYVNIIDRSSGDTVLIGAVGGHVEIVRAL 290  
 506 NAATDLYTPLHIAKEGEEVAALIMDHGTDKTLTKKFTPLHIAKYNENLVAAKSL 565  
 291 OKYADIDIRGODNKTALYMAVEKGNATVNRDILQCNPTETCTDGEPLIKAKRMNIE 350  
 566 EKGTVYDEGKNQVPLHVAHYNNDKVALILLEGASAHAAANGYTPHIAAKNQMD 625  
 351 VELLLDGKAVSAVDKGDTPHIAIRGSRKLAELLRN----- 391  
 626 IASTLLHYKANANESKAGFTPLHIAQEGHREMAALLIENGAVGAQARNGLTPMHLCA 685  
 392 -----PKGRLLYPRNKAGFTPNINDCHOSILTOIFGARHLSFTEDGMLG 440  
 666 QEDRVSAVEELVKEAALDPKTKAGYTPHIAVACH-----FG-----QIMNVR 727  
 441 YDLYSSALADILSEPTMOPICVGLIYAQWGSKSFLLKLEDEKKTAGQOIEPLPQFSW 500  
 728 FLIEHGAARVITRASYR-----LHQAAQGHNSYVR----- 760  
 501 LIVFTLLCGGLGILAFVYHPULGIAVLSLFTALLYFTFYVFGRRREGESMNAWV 560  
 761 -----YLEHGASPNVHTSGOTPLS----- 781  
 561 LSTRLAHIGYELLKLMFVNPPELPEOTTKALPVRF-----LFTDYNRLSSV 609  
 782 -----IAERLGYVYVEALKTTETTYVITETTYTEERIKKPPONPEANMETMSD---SED 833  
 610 GGSTSLAEMITLSDACEREFGFL-ATRLFRVETEDTQGGK----- 650  
 834 EGEDNQITANAHADFESLTKGLHDSYGVALIHATEPTLSRSEVEVGSTGDALIRKA 893  
 651 -----KWKTCCLPSFYIPLFT-----IGCIISGTTLAI 681  
 894 QHEPITAMADPSLDASLPNVNITMRTMOPSLISPMVAVARGAMGCRHSGRITIPP 953  
 682 RVDKHLTVANAVALISASVGLAFVLCRTMOWLDLNLN-----SOKRKLHNAASKLHL 737  
 954 RKAQOPTRV-----TCR-YLCKDKLAHPPLISEGAL---ASRLLEM 991  
 738 KSEG--FMKYLKCEVELMARKATIDFTQNGTRLVYI--DG-----LDACRODKV 785  
 992 APHGAKFLGPILEVPHFASL-----RGREHEIYILASDGOHKKHEQLATE-DAV 1042  
 786 LQMLD-----TVRLKSGPPIAFASDPHIIKAINQNLNLYVLOS 827  
 1043 QEVYNESFDAEELSOLDLHTSRITRLITLNDPFP--YFA-----VTVRQVEHCVGPEG 1095  
 828 NINCHDYMRNIV-HLPEVLNRSGLSNARKFLVTSATNGDVCSDTGTGQDADARRYSQNS 886

1096 GV-----ILSSVPHVQALFPDGSILT---KTIKYSVAQAPVQEIYVRLH----- 1137  
 887 LGEMTKLGSKTALNRAPTY--RRROMRTTTRQMSFDLTK--LVTEDWFSDISPQNR 941  
 1138 -----GNVAVASPIYTVPRRRKFKPITLTCIPLRQSSKNKMLQYSGQPEQPTIR 1190  
 942 RL-----NIVSYTGRLLRANQISEN-----W-----DLASWI-N 971  
 1191 LKCSKTGSSPAQWEDITCTTQLTFTGEDVSFTTVASARFWMDCOTPRDAARMAQEVYN 1250  
 972 LTEQWPRYSMLIYEEETGIPDOMTL-----KIT--YERISKNIPTTKDEPL- 1019  
 1251 EATAVPMARFLI-FARTTPAGQLRFLCMTDREDKTKLEKQERTIE-IASKDVEVLIS 1308  
 1020 -----LEIDGDIRNEFEVLESSRTPVLVARD--VKVFLPCTVN--LDPLR-----EITA 1064  
 1309 GRHOFLEFSGNL-----LPITKSGDLSLYFLPQENRLAPMVKIRHTDNETA 1358  
 1065 DVNARQOISIGLAVPLPLHGGPRRABGYSQPSVCSSTISFNCPFAGVVSQPHSS 1124  
 1359 DGR-----IVMKKEPKLAENLPOTP-----VCTLAITLPEYTG--PEPMVS 1399  
 1125 ---YYS--GMTGPHFYNRSGPAPGP--VVLNSLVNDVACEKTK-----Q 1165  
 1400 KILFYSASLTERKYVAGFHEHTAPDMLPLAHVALLIGAMHRLARALEVPDIDIRQVRHQ 1459  
 1166 IEGDQSMLEPOYCTTKKANIN--GRAVLAQCNIDELKEMNM-----F 1207  
 1460 LVGLFAVYTLIRWIFLKEQATPVALRSALORIGRDVAREMDRAKLDLGGTPVSHIS 1519  
 1208 GDMHFRSTVYLEM-----RNAE--SHVVEDPRFLSSSSG-PAPHEPARARSHNLT 1257  
 1520 GBSITSLILEVAGRRRAEYVYMAQRIADPEPQOVGYVGTQEDPEEPQOSFHEE 1579  
 1258 PATELSSQPPYTLNFSPEELNLTGLDEGAPRHSNLSMOSQTRTPSLSSINQSDSIETS 1317  
 1580 EEVAVS-----EIRTVVRE--RHVHSENGPIYEBERITTYEDDVAVNEE 1624  
 1318 KLUDY-----QAEYRAYREYIAQMSQLEGGCSTIISGRSSPHSYTYGSSGGS 1370  
 1625 EYVDKIVPLNEEQEADMRVREVENFQOE-----TSKEGTFCQQTTHREKDDGGS 1679  
 1371 IHSNLEQEKDSEKPP--DDGRKSFLMRKRDVIDYSSSGYSTNDASP-LDPIFEED 1424  
 1680 LKTTM-----KQSHVQIFPDGGETS-----ANETGLSSDADQIMPTTKED 1722

RESULT 6  
 US-09-577-034-33  
 ; Sequence 33, Application US/09557034  
 ; Patent No. 6365569  
 ; GENERAL INFORMATION:  
 APPLICANT: Tang, Liang  
 Blehm, E. Scot  
 TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
 PROTEINS, NUCLEIC ACID MOLECULES, AND  
 USES THEREOF  
 NUMBER OF SEQUENCES: 171  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Carol Talkington Verser, Ph.D.  
 Heskia Corporation  
 STREET: 1825 Sharp Point Drive  
 CITY: Fort Collins  
 STATE: Colorado  
 COUNTRY: USA  
 ZIP: 80525  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: Windows 95  
 SOFTWARE: Wordperfect for Windows, Version 7.0  
 CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/09/557, 034  
 FILING DATE: 21-Apr-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/065,474  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Verser, Carol Talkington  
 REGISTRATION NUMBER: 37,459  
 REFERENCE/DOCKET NUMBER: HW-5-C1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 970/493-7272  
 TELEFAX: 970/484-9505  
 INFORMATION FOR SEQ ID NO: 33:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1745 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
 US-09-557-034-33

Query Match 5.4%; Score 480.5; DB 4; Length 1745;  
 Best Local Similarity 20.3%; Pred. No. 5.6e-32;  
 Matches 353; Conservative 234; Mismatches 602; Indels 549; Gaps 67;

14 EENIPALKALKECKVDERNECGOTPLMAEOGLYEKELIKGANCLEDDN-- 71  
 207 KDDYTKATLLQNEHNSDVTSKSGFTPLHAHYGENVAOALLEKGAVANNQARHNIS 266  
 72 -----W----- -TALISAKSEGHVHYEEELKCG----- 94  
 267 PLVATKMGRTNMVSLLAHGAVIDCRTLRLTLPLHCASSGSHQVVDLLEKCAPISAK 326  
 95 -----VNEH----- -RDMGWTALMW 110  
 327 TKNGPLAHMAAOYVDVDTYPLPLVHAHCGHVRVAKLLDRNADPNARALNFTPLHI 386  
 111 ACYGRDIDVVELLSHGANSVYGLQSYPIYMAAGRGHADIYHLLQNGAKVNCSDKY 170  
 387 ACKNRKRIYELLKRYHAIEAT-TEGSLPLVHAAPFAGALNIVYLLQOGANADVATVR 445  
 171 GTPLVMAARKGHECVKHLAMGADVDEGANSMTALIVAKGYTOSYVEILKRPNV 230  
 446 GETPLHAAKANDIYVLRNGADVAARELOTPPLHISRLGNDIYLLQANASP 505  
 231 NLTDGNTALMASKEGHEIYODLDAGTYVNIIPDSGPTVLIGAVRGHVEIVALL 290  
 506 NAATRDLYTPLHIAKEGOEEVAAIIMDHGTDKTLTKKFTPLHAKYGNLPAKSL 565  
 291 OKYADIDIRGODNTALVWAVEKGMATVNRDILQCNPTLEICTDGEPLIKATKMRNIE 350  
 566 EKGTPVDIEGKNQVPLVAAHYNDKVALLLLEGGASAAHAAKAGTTPHIAAKKNOMD 625  
 351 VVELLDGAKVSAVDKKGDTPLHAIAGRSKRLAEELLRN----- 391  
 626 IASTLHAKANNAASKAGFTPLHIAQEGHREMAALLIENGAKVGAQARNGLPMLCA 685  
 332 -----PKGRLLYRPNKAGETPYNIDCSHOKSILQIIGARILSTPEIDGMDLG 440  
 686 QEDRVSAEELVKEAAIDPKTKAGYTPLHVACH-----FG-----QINVR 727  
 441 YDLSSALADILSEPTMQPICVGLYAQMSGKSLKLEDEMKTFAQOIEPLFQPSW 500  
 728 FLIEHGAARYVITRSTYR-----LHOAAQOCHNSVVR----- 760  
 501 LTVPLTLLCGIGLGLFAFYVHPNLGIANSLSFALLIIFPLVIYEGRRGESWMAV 560  
 761 -----YLIEHGA SPNVHTSGCPPLS----- 781  
 561 LSTRLAHIGYLELLKLMEFVNPPLPBOITKALVVR-----LFTDYNRLSSV 609  
 782 ----IAERLGYVVAELKTTTETVITETVTEERYKPPONPEAMNETMSD-----SED 833

610 GGETSLAMIAINTLSDACERERGF--ATRLFRVETEDYOGKK----- 650  
 834 EEDNQITANAHADPSLSLKLGHDSGVHLIATETPLSRPVEGTDDDLALIRKA 893  
 651 -----KMKTCCLPSEVIFLPI-----ICGIIIGTLLAIF 681  
 894 QHEPITTMADPSLSDASLPDNTIMRTMQPSFLISFVVDARGAMRCGRSHGVAILPP 953  
 682 RVDPRHLVYNAVLISTASYVGLAFVLCNRTMQVDSILN-----SQKRLNNAASKLHL 737  
 954 RKAQPTPRV-----TCR--YLGRDKLAHPPPLSEGEAL--ASRILEM 991  
 738 KSEG--FMKVLKCEVELAMAKTIDSFONOTRLVIT--DG-----LDACEODKV 785  
 992 APHGAKEFLGVLLEVPHPASL-----RGREIVILRSDGGMKHEHLEATE-DAY 1042  
 786 LQMLD-----TVRLFSGKPTAIPASDPHIIKAINQNLNSVLD 827  
 1043 QEVLNMFDAEELSOLDLHTSRITRLINDEPM--YFA-----VTVRQVEVHCVGEG 1095  
 828 NINGHYMKNIV-HLPVFLNSRGLSMARKELVTSATNGDVPDGTGIEDADRRVSONS 886  
 1096 GV-----ILSSVPHVOAIFPDGSL--KTIVSOAOPVPOEIVTRLH----- 1137  
 887 IGEHTKIGSKTALNRDLY--RRQMQRTITRQMSFDLTK--LVTEDMFSDISPQWR 941  
 1138 -----GNRVASPLVYVEPRRRKFKITLICILPQSSNKGMLTOYSGQGPQEPPLR 1190  
 942 RLL-----NIVSYGRHLRANOISRN-----W-----DRLSWT-N 971  
 1191 LILCSKTGSSPAQWEDIEGTTLQTLFTGEDVSEFTTVYSAFWMDCQTPDAARMQOEYV 1250  
 972 LTEQWPTKSWLILYEETEGIDPQDL-----KTI--YERISKNIPTTKDEPL- 1019  
 1251 EATVPPMAKEL-FARKTPPAEQRLRCMDMDREDKLEQERTE-IASKOYEVL 1308  
 1020 -----LEIDGIRNFVEFLSSRTPLVARD--KVYFLPCTVN--LDPLR-----ETIA 1064  
 1309 GRHQFLEFSGNL-----LPIKSGDQLSLYFLPQENRLAFMRKIRHTHDETA 1358  
 1065 DVRAARQISITGLATPLPLHSGPPRABSGYQSPSSCSISFNPFAGGVVPOPHSS 1124  
 1359 DGR-----IVMKBEKLAEMLPQTP-----VCLATLDEPYTG--PEPMVS 1399  
 1125 ----YIS--GMTGPHRPNYNGSGPAPG--VVLNSLWDAVCEKIK-----Q 1165  
 1400 KILFYSASITTEKYVGAHFHTAEPDNLPLAHVALLIGADHRLARALEVPDIDIRQVRH 1459  
 1166 IEGIDQMLPOYCTTIKKANIN-----GRVLAQCNIDELKEMNM-----F 1207  
 1460 LVGLEAVTILRIWFLKKEQATPVALSALORIGRDVVAEMRAKLDLGEPTVSHIS 1519  
 1208 GDMHLEFSTYLEM-----RNAE-----SHVVEPDRFLESSESSG-PAPHGEPARASHNEL 1257  
 1520 GPSITLSTLLEAAGDRRRRAHEVYMAOQRLAOEFPFOQVYKATPGDPEPKQSHHEE 1579  
 1258 PHELSQPTTYTLNFSFEELNTLGLDGAHRHSLWSQOTRRTPSLSSINSOSSIEIS 1317  
 1580 EVAVS-----EIRTVRTE--RHVHDSBNGPIVERETITTYEDVAVNEE 1624  
 1318 KLNDK-----QAEYRDARREYIAQMSLEGPGSTISGSSPHSTYMOQSSSGS 1370  
 1625 EIVDKIYPLNEEEDKMDRMVREKMNFEQOE-----TSKEGTFGCQTHEKEKDDGGS 1679  
 1371 IHSNLEDEKGDSPKPR--DDGRKSFMLKRGVDIVDSGSGVSTNDASP-LDPIEDED 1424  
 1680 LKTYM-----KDSHVROIIFPDGETS-----ANETGLSSGADMTIMTITKED 1722

RESULT 7  
 US-09-082-059-2  
 ; Sequence 2, Application US/09082059A



Db 1109 W 1109

## RESULT 9

US-09-196-387-2  
Sequence 2, Application US/09196387

Patent No. 6277613

GENERAL INFORMATION:

APPLICANT: de Lange, Titia

APPLICANT: Smith, Susan

TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

TITLE OF INVENTION: OF USE THEREOF

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauder &amp; Jackson

STREET: 411 Hackensack Avenue, 4th Floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/196,387

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/095,225

FILING DATE: June 10, 1998

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-230 CIP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1327 amino acids

TYPE: amino acid

SPRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-09-196-387-2

Query Match

Best Local Similarity 22.3%; Pred. No. 4.4e-23;

Matches 160; Conservative 89; Mismatches 220; Indels 247; Gaps 17;

QY 14 EENENIPALKLEKCKDVDERNECGOTPLMAAEOGNEIYELIKNGANCNLEDDJNMT 73  
Db 348 EEKIMALLPLPLNVNCHASDGRK---STPLHLAAGYNRVYIOLLLHOGADVHAKKGLV 404  
QY 74 ALISASKBHVHIVVELLKCGVNLHRDGGFTALMMACYKRTYVELLSHGANSPT- 132  
Db 405 PLHMAGSYGHEVTELLKNGACVNAIMDLOFTPLHKAASKRVEVCSLLSHGADPTLV 464  
QY 133 -----T 133  
Db 465 NCHGKSAVDMAPTPELRELUYEFKSHSLQAAAREADLAKVKTALLETINFKOPSHET 524  
QY 134 GLQYSV-----YPIIWAARGHADIVHLLONGAKV 164  
Db 525 ALHCAVASLHPRKQYTELLLRKGANVNEKNRDFMPLHVAABRAHNDVMEVLHKGAKM 584  
QY 165 NCSDKYGTPIVWAARKGHLKGVKLLHMGADVDGAGNSMTA----- 207  
Db 585 NALDTLGGTALHRAALAGHLQYCRLLSLYSGDPSIISLOGFTAAOMGNBAVOOIISESTP 644

QY 208 -----LIYAVKGGYQSVKEILKRNPNVNLTDKDS--NTALMASKEGTEIYODL 256  
Db 645 IRTSDVYRLLEKSKAGDEYKQDLS-SQNNCRDLBGRHSTPLHFAAGYKRVSEVL 703  
QY 257 LDAGTYVNIPIRSGDFTVLIGAVRGHVEIVRALLQYADIDIRGODNTALYWAYEK-- 314  
Db 704 LHHGADVHAKKGLVPLHNNACSYGHYEVAELIVRGASVNVADLMKFTPLHAAKGY 763  
QY 315 -----NATVVRILQ-----C 325  
Db 764 EICKLLKGGADFTKKNRDNPTLVKSGDDIDLLGDAALLDAKGLARVQKLC 823  
QY 326 NPDEICTKD---GETPLKATMRNIEVELLDGAVSAVDKKGTPPLH----- 374  
Db 824 TPEINCRDTPQGNSTPLHAAQYNNLEVAEYLLHGGADVNNQDKGLPLHNAASGYH 883  
QY 375 -----IAIRGRSKLAELLRNPKDGRLLYRPNKAGTP 408  
Db 884 DIAALLIKYNTCVNATDKNAFTPLHBAOKGR-QLCALLHAGADPTM---KNQGGTP 939  
QY 409 YN-----IDCSHOKSILT-----QIFGARHLSPTETDGMGLGYDYSALADI 451  
Db 940 LDLATADDIRALLIDAMPPLPCTCRKPOATVVSASLISPASTP-----CLSAASIDN 994  
QY 452 LSEPTQPIPCVGLYQWGSKSFLLK-----LEDEKRTFAGQIIEPLFQ 497  
Db 995 LTGPIAB--LAVGASAGADGAAGTERKEGEVAGLDNMNISQFLKSLGLEHLRDI 1048

RESULT 10  
US-09-031-485-28  
Sequence 28, Application US/09031485  
Patent No. 5824306  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
APPLICANT: Blehm, E. Scott  
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 85  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESS: Hesk Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/031,485  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/847,429  
FILING DATE: 24-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

Query Match 4.2%; Score 374; DB 2; Length 348;  
Best Local Similarity 33.7%; Pred. No. 5.5e-24;  
Matches 112; Conservative 48; Mismatches 134; Indels 38; Gaps 4

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

TYPE: amino acid  
TOPOLOGY: linear

```

? INFORMATION FOR SEQ ID NO: 10:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 949 amino acids
?

```

TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-196-387-10

Query Match 4.2%; Score 370.5; DB 4; Length 949;  
 Best Local Similarity 24.1%; Pred. No. 7,7e-23;

Matches 139; Conservative 62; Mismatches 174; Indels 201; Gaps 11;

```

QY 14 EENIPALKEKDKDERNECQOTPLMAEAGNEIYVELKNGANCNLEDLDWMT 73
DB 348 EKKMALPLPLVNVNCHASDGRK---STPLHLAAGNRRRIYOLLQHGADVHARDKGLV 404
QY 74 AIISSAKBGHVIIVELLKCGVNLHRDMSGCTALMAQYGRDVEYLLSHGANSV- 132
DB 405 PLHNACSYGHEVETELLKHGACVNAAMDLOFTPLHEAASKNRVEVCSLLSHGADPTLV 464
QY 133 -----T 133
DB 465 NCHGSAVDMAPTPELRERLYEFKSHSLQAAAREADLAKYKTLALEIINEKOPQSHET 524
QY 134 GLQYV-----YPIIWAAGRHADIYHLLONGAKV 164
DB 525 ALHCAVASLHPKRGQVTELLRKGANVEKNDFPTPLHVAERAHNDVMEYLKHKAKM 584
QY 165 NCSKXGTPPLVMAARKGHECVKHLAMGADVDOEGANSMTA----- 207
DB 585 NALDPLGQTHLRALAGHLQTCRLLSYGSDPSIISLQGFYAAQMGNEAVQQLSESTP 644
QY 208 -----LIIAIVGGYTSYKELKRNPNVNLTDKQ--NTALMISKSGHEIYQDL 256
DB 645 IITSVDYRLLEASAGDLETYKQLCS--SQVNCDELRHSTPHFAAGYNRVSVETL 703
QY 257 LDAGTYVNIPIRSGDTVLIGAVRGHVEIVALLQYADIDIRGDNKTALYWAVERG-- 314
DB 704 LHHGADVNAKDKGLVPLHNACSYGHEVALLVHNGASVNAVADLMKTPPLHEAAKGY 763
QY 315 -----NATVARDLIQ-----C 325
DB 764 EICKLLKHGADPTKKNBDGNTPLDLVKEGDTIDQLKGDALDLAAKKGCLARVOKLC 823
QY 326 NEDTETCTKD--GETPLIKATKRNIEVEVELLDKAGKYSANDKGDPTLHIAIR----- 378
DB 824 TPEINNCADYGRNSTPLHLAAGYNNLEVAEYLLHEGADVNAODKGLIPLHMAASYGCC 883
QY 379 -GRSRKLAELLRNPKDGRLLYRPNAKGETRYNIDC 413
DB 884 IARVOKLC-----TPEINNC 898

```

RESULT 15  
 US-09-350-982C-5  
 Sequence 5, Application US/09350982C  
 Patent No. 6455290

GENERAL INFORMATION:  
 APPLICANT: Berthelsen, Jens  
 APPLICANT: Toma, Salvatore  
 APPLICANT: Isacchi, Antonella  
 TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods Relat  
 FILE REFERENCE: PHM-0043  
 CURRENT APPLICATION NUMBER: US/09/350,982C  
 NUMBER OF SEQ ID NOS: 10  
 SOFTWARE: Patent version 3.0  
 SEQ ID NO 5  
 LENGTH: 1166  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 NAME/KEY: misc\_feature  
 OTHER INFORMATION: xaa is any amino acid  
 NAME/KEY: misc\_feature

LOCATION: (1102)..(1102)  
 OTHER INFORMATION: n is any nucleic acid  
 NAME/KEY: misc\_feature  
 LOCATION: (2650)..(2650)  
 OTHER INFORMATION: n is any nucleic acid  
 US-09-350-982C-5

Query Match 4.1%; Score 363; DB 4; Length 1166;  
 Best Local Similarity 22.9%; Pred. No. 5.2e-22;

Matches 141; Conservative 65; Mismatches 181; Indels 228; Gaps 12;

```

QY 19 PALKALLEKCD-----VDERNECGQ--TPMLTAEQNEIYVELIKNGA 62
DB 23 PAARELFACRNDVERKRLVPEKYNSHDTAGKSTPLHLAAGGRDVEYELLQNGA 82
QY 63 NCNLEDNDNTALISAKBGHVIIVELLKCGVNLHRDMSGCTALMAQYGRDVEYEL 122
DB 83 NVQARDDGGLIPLHNACSYGHEVALLVHNGADVPARNDWNTPLHEAALIKIDVCIY 142
QY 123 LLSHGANSV-----TGLQY 137
DB 143 LQHGAEPTIRNTDGRALDLADPSAKAVLTGEYKKDELLESARSGNEERKMAALLTPLV 202
QY 138 SVT-----PIIWAAGRHADIYHLLONGAKVNCSDKYGTPPLVMAARKGHE----- 185
DB 203 NCHASDGRKSTPLHLAAGYNRKIVYOLLQHGADVHAKDKGDLVPLHNACSYGHEVTEL 262
QY 186 -----CV----- 187
DB 263 LVKHGACVNAAMDLOFTPLHEAASKNRVEVCSLLSYGADPTLLNCHNSAIDLAPTPOL 322
QY 188 -----KH----- 189
DB 323 KERLAEYFKGSHSLQAAAREADYTRIKHLSLEWVNFKHROTHTAXHCAASPYPRKQI 382
QY 190 ---LLMGADVDOEGANSMTALVAVKGGTQSVKELKRNPNVNLTDKGTALMIAASK 246
DB 383 CELLRKGAXINKEKTEFTPLHVASERKANDVEYVNHKAVNALMDIGOTSLHRAV 442
QY 247 EGH-----TEIVODLDAGTYVNIPIRSGDTVLIGAV 278
DB 443 CGHLQTCRLLSYGCDPNIIISLOGFTALOMGNENVOQLQEG--ISLGNSEADROLLEA 500
QY 279 RGHVEIVRAL--LQYADIDIRGDNKTALYWAVERGNATYVRDILQCNPTETICTKG 336
DB 501 KAGDVEYVKKLCTVQSVNCRDIEGROS--TPLHFAAGYNRVSVETYLQHGADVNAKDKKX 559
QY 337 ETPLIKATKRNIEVEVELLDKAGKYSANDKGDPTLHIAINGRSKLAELLRNPKGR 396
DB 560 LVPLHNACSYGHEVALLVKGAVVNAVADLMKTPPLHEAAKKGKEYICKLILQHGADPT 619
QY 397 LLYRPNAKGETRYNI 411
DB 620 ---KKNRDGTPLDL 631

```

Search completed: July 1, 2003, 14:41:04  
 Job time : 39.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 16 Seconds  
(Without alignments)  
4445.744 Million cell updates/sec

Title: US-10-021-571-2

Perfect score: 8853  
Sequence: 1 MSVLISQSYNYVEENRPA.....ELHAASSESTGCEERESTL 1715

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556	6.3	1880	1 ANK1_HUMAN	P16157 homo sapien
2	550	6.2	3924	1 ANK2_HUMAN	Q01484 homo sapien
3	527.5	6.0	1862	1 ANK1_MOUSE	Q02357 mus musculu
4	526	5.9	4377	1 ANK3_HUMAN	Q12955 homo sapien
5	468	5.3	768	1 VR23_HUMAN	Q91177 homo sapien
6	431.5	4.9	1059	1 VR39_HUMAN	Q15064 homo sapien
7	414.5	4.7	1431	1 DAPK_HUMAN	P53353 homo sapien
8	395.5	4.5	1327	1 TNK1_HUMAN	Q95271 homo sapien
9	384.5	4.3	1327	1 ASB2_HUMAN	Q96Q27 homo sapien
10	380.5	4.3	1166	1 TNK2_HUMAN	Q912K2 homo sapien
11	378.5	4.3	323	1 ANKH_CHRYT	Q06527 chromatium
12	360	4.1	832	1 ANK3_HUMAN	P57078 homo sapien
13	357.5	4.0	741	1 RNSA_HUMAN	Q05823 homo sapien
14	356	4.0	735	1 RNSA_MOUSE	Q05921 mus musculu
15	349.5	3.9	692	1 ANR6_HUMAN	Q91294 homo sapien
16	335	3.8	1401	1 LATR1_MOUSE	P26311 latrodectus
17	331	3.7	525	1 ASB3_MOUSE	Q94V72 mus musculu
18	329	3.7	583	1 ASB5_MOUSE	Q8V86 mus musculu
19	328.5	3.7	518	1 ASB3_HUMAN	Q95755 homo sapien
20	328.5	3.7	747	1 V222_FOPPV	Q91513 fowlpox vir
21	310.5	3.5	542	1 V155_FOPPV	Q915A7 fowlpox vir
22	309	3.5	668	1 V244_FOPPV	Q914Z6 fowlpox vir
23	308	3.5	596	1 V024_FOPPV	Q915B7 fowlpox vir
24	307	3.5	436	1 V245_FOPPV	Q914Z5 fowlpox vir
25	288	3.3	523	1 V228_FOPPV	Q91507 fowlpox vir
26	288	3.3	1083	1 V112_YEAST	P40480 saccharomyc
27	271	3.1	592	1 V246_FOPPV	Q914Z4 fowlpox vir
28	270	3.0	603	1 V162_FOPPV	Q91569 fowlpox vir
29	265.5	3.0	434	1 AS15_HUMAN	Q8VXK1 homo sapien
30	264.5	3.0	656	1 FEHL1_MOUSE	P17221 caenorhadi
31	260	2.9	433	1 AS14_MOUSE	Q8V87 mus musculu
32	254.5	2.9	776	1 ANR5_HUMAN	Q91N02 homo sapien
33	251	2.8	429	1 AS10_HUMAN	Q8VX13 homo sapien

34	245.5	2.8	578	1 V022_FOPPV	Q915H9 fowlpox vir
35	245	2.8	278	1 AS13_HUMAN	Q8VXK3 homo sapien
36	242	2.7	434	1 V023_FOPPV	Q915B8 fowlpox vir
37	238.5	2.7	333	1 ANR2_HUMAN	Q94ZV1 homo sapien
38	237.5	2.7	1066	1 NDC2_MOUSE	Q092V1 neurospora
39	236.5	2.7	328	1 ANR2_MOUSE	Q94V06 mus musculu
40	235.5	2.7	2703	1 NDC2_MOUSE	P07207 drosophila
41	235	2.7	231	1 NSDC_MOUSE	Q94ZK3 rattus norv
42	235	2.7	642	1 V2A2_SCHPO	Q09701 schizosach
43	234	2.6	437	1 V014_FOPPV	Q91517 fowlpox vir
44	230.5	2.6	461	1 V218_FOPPV	Q91517 fowlpox vir
45	230	2.6	226	1 PSDA_HUMAN	Q75832 homo sapien

## ALIGNMENTS

## RESULT 1

ID ANK1\_HUMAN STANDARD: PRT: 1880 AA.

AC P16157; 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).

GN ANK1 OR ANK.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_Taxid=9606;

RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.

RC TISSUE=Hematopoietic;

RA MEDLINE=90158830; PubMed=2137557;

RA Lux S.E., John K.W., Bennett V.;

RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins.";

RT Nature 344:36-42(1990).

RL [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=9015730; PubMed=1689849;

RA Lambert S., Yu H., Pichal J.T., Lawler J., Ruff P., Speicher D.,

RA Cheung M.C., Kan Y.W., Palek J.;

RT "cDNA sequence for human erythrocyte ankyrin.";

RT Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).

RL [3]

RP VARIANT HS ILE-462.

RA MEDLINE=96225450; PubMed=8640229;

RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,

RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,

RA Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;

RT "Ankyrin-1 mutations are a major cause of dominant and recessive hereditary spherocytosis.";

RT Nat. Genet. 13:214-218(1996).

RL [1] FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN GPB5, AND TO THE NA-K ATPASE. TO THE LYMPOCYTE MEMBRANE PROTEIN GPB5, AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN. ERYTHROCYTE ANKRYNS ALSO LINK SPECTRIN (BETA CHAIN) TO THE CYTOSKELETAL DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN; THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.

CC [1] SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC PLASMA MEMBRANE.

CC [1] ALTERNATIVE PRODUCTS: At least 3 isoforms; 1/2.1 (shown here), 2/2.2 and 3 are produced by alternative splicing.

CC [1] PPM: REGULATED BY PHOSPHORYLATION.

CC [1] PM: ACYLATED BY PALMITIC ACID GROUP(S).

CC [1] DISEASE: Defects in ANK1 are the cause of dominant and recessive hereditary spherocytosis (HS).

CC [1] SIMILARITY: CONTAINS 23 ANK REPEATS.

CC [1] SIMILARITY: CONTAINS 1 DENTH DOMAIN.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; X16609; CAA34610.1; -  
DR EMBL; M28880; AAA51732.1; -  
DR PIR; S08275; SJHUK.  
DR PIR; A35049; A35049.  
DR HSSP; 000420; 1ANC.  
DR Genew; HGNC:492; ANK1.  
DR MIM; 162900; -  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR000906; ZUS.  
DR Pfam; PF00023; ank; 24.  
DR Pfam; PF00531; death; 1.  
DR Pfam; PF00791; ZUS; 1.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 22.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00218; ZUS; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 20.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
DR Cytoskeleton; Alternative splicing; Repeat; ANK repeat;  
KW Phosphorylation; Lipoprotein; Disease mutation; Elliptocytosis;  
KW Polymorphism.  
FT INTL MET 0 0  
FT DOMAIN 1 826  
FT 89 KDA DOMAIN (ANTION EXCHANGE PROTEIN  
FT DOMAIN 827 1381 BINDING DOMAIN).  
FT 62 KDA DOMAIN (SPECTRIN BINDING  
FT DOMAIN 1382 1880 DOMAIN).  
FT 55 KDA REGULATORY DOMAIN (REGULATES  
FT THE BINDING OF ANKYRIN TO SPECTRIN  
FT AND THE BAND 3 PROTEIN).  
FT REPEAT 43 72 ANK 1.  
FT REPEAT 76 105 ANK 2.  
FT REPEAT 109 138 ANK 3.  
FT REPEAT 142 171 ANK 4.  
FT REPEAT 173 200 ANK 5.  
FT REPEAT 204 233 ANK 6.  
FT REPEAT 237 266 ANK 7.  
FT REPEAT 270 299 ANK 8.  
FT REPEAT 303 332 ANK 9.  
FT REPEAT 336 365 ANK 10.  
FT REPEAT 369 398 ANK 11.  
FT REPEAT 402 431 ANK 12.  
FT REPEAT 435 464 ANK 13.  
FT REPEAT 468 497 ANK 14.  
FT REPEAT 501 530 ANK 15.  
FT REPEAT 534 563 ANK 16.  
FT REPEAT 567 596 ANK 17.  
FT REPEAT 600 629 ANK 18.  
FT REPEAT 633 662 ANK 19.  
FT REPEAT 666 695 ANK 20.  
FT REPEAT 699 728 ANK 21.  
FT REPEAT 732 761 ANK 22.  
FT REPEAT 765 794 ANK 23.  
FT DOMAIN 1402 1486 DEATH.  
FT VARSPLIC 1512 1873 MISSING (IN ISOFORM 2).  
FT VARSPLIC 1874 1874 H -> D (IN ISOFORM 2).  
FT VARSPLIC 1849 1880 GLOPDLIEGRKAQIVKASIKRKQ (IN ISOFORM  
FT 3).  
FT R -> T.  
FT /FTid=VAR\_000595.  
FT V -> I (IN HS).  
FT /FTid=VAR\_000596.  
FT R -> H (IN BRUEGGEN).

FT FT  
FT VARIANT 749 749 /FTid=VAR\_000597.  
FT V -> A.  
FT /FTid=VAR\_000598.  
FT VARIANT 844 844 D -> E.  
FT /FTid=VAR\_000599.  
FT VARIANT 1285 1285 E -> D.  
FT /FTid=VAR\_000601.  
FT VARIANT 1391 1391 S -> T.  
FT /FTid=VAR\_000600.  
FT VARIANT 1591 1591 D -> N (IN DUESSELDORF).  
FT /FTid=VAR\_000602.  
FT VARIANT 1698 1698 R -> D.  
FT /FTid=VAR\_000603.  
FT CONFLICT 229 A -> S (IN REF. 2).  
FT CONFLICT 1545 V -> I (IN REF. 2).  
SQ SEQUENCE 1880 AA; 206145 MW; 1C5F5E7EED1CD448 CRC64;

Query Match 6.3%; Score 556; DB 1; Length 1880;  
Best Local Similarity 20.6%; Pred. No. 1.2e-22;  
Matches 395; Conservative 270; Mismatches 653; Indels 620; Gaps 75;

QY 21 LKALLEKCKVDNERNCGQTPPLMAEGCNVEIVKELKNGANCNLEDNNTALLISAK 80  
DB 93 VRELVTYGANVNAQSGKFTPLMAQENHLEVKFLKNGANOVATEDGFTPLVALQ 152  
QY 81 EGH----- 83  
DB 153 OGHEENVAMHLYNGTKGVRLPALHIAARNDGTRPAVLVLDNDPNDVLSTGTPLHIA 212  
QY 84 -----IHYBELKKSASLEHRDMSGWTLMAACRYGRDVEYLLSHGANSVTGLQYS 138  
DB 213 AHENLVNAQLLNRGASVFTPONGITPLHIASRGVIVVRLLDGQAQIE-TKTDE 271  
QY 139 VYPTIAGRGHADIYVHLLONGAKVCSDKGTGTPPLVAARKHLECYKHLANGAVD 198  
DB 272 LTPHLCAARNGHVRISIIIDHGAPLOATKGLSPIHMAOGDHLDVRLLOYDAEID 331  
QY 199 -----QEQA-----NSMTALVAVKGQYQSYKEIK 225  
DB 332 DITLDLTPPLHVAHCGHHRVAKVLLDKGAKPNSNALNGFTPLHACKNHVRMELLK 391  
QY 226 RNNVNLVTRKQDNTALMTSKEGHEIYVDLDAGTYVNIIPRSDDYILGAVRGHVEI 285  
DB 392 TGAASIDAVTESGILTPPLHVASFMGHLPYKNLLDORASPVSVKVEYTPHMAARAGHREV 451  
QY 286 VVALLOKYADIDIRGQDNKTAUYWAVEKGNATVNDILOCNPDTEICTKDEFTPIKATK 345  
DB 452 AKYLLONKAKVNAKAKDDOTPLHCAARIGHTIMVKKLLENNANPMLATTAGTPLHIAAR 511  
QY 346 MRNIEVVELLDKGAQVSAVDRKGDTPPLHVALRGSRRLAELLNRPKDGRLLYRPNKAG 405  
DB 512 EGHVEVETLALLLEKASQACMTKKGFTPLHVAKYGVVAELLE--RDA---HPNAG 565  
QY 406 E---TPYNIDCSHQ-----KSTL-----TPIGAR----- 427  
DB 566 KNGLTPPLHVAHNNLDIYKLLPRGSGSPHAWNGYTPPLHIAKONOVEVARSLLOYG 625  
QY 428 -----HLSPETDGMGLGYDLYSSALADILSEPTMOP-----PIC-- 462  
DB 626 SANASVQGVPLPLHAAEGHAEMVALLSKQANLNCGKSLTPPLHVAOEGHVPADV 685  
QY 463 -----VGLYAOMGSGK--SFLIKLIED-EMKTFAGQOTEPLPORSWL 501  
DB 686 LKHGVMVDATTRMGYTPPLHVASHGNIKLYKFLLOHODAVNAKTKLGY--SPHQAQQ 743  
QY 502 --IVFTLLTLGGGLVAFAPVD-----TNAATAISLFTALITYFIYIYVPGRRGE 553  
DB 744 GTDITVTLKNG-----ASPNEVSDGTTPLHAKRIGIYSIVDVLKVV----- 788  
QY 554 SNWMAWALSTRARHIGYLELLEKLMFVNPPELPEOTKALPVRFLETDYNRLS-SVGE 612  
DB 789 TDETSFVLVSKHR-MSFPETVDEILDVSEDBGEELIS-----FRAERDRSDYDEE 839

```

QY 613 TSLAEMIAITLSDACEREFGLATRLR-----VPRTEE-SQCKKKKKTKCCLSPV 662
DB 840 KELLDPVPKLDQVE-----SPAIPRICAMPETVIVRSEEDQSKKEDESLIPS-- 891
QY 663 IFLEIVGCIAGITLTAIFRVDPKHLTVNALISIASVGLAFVLCRTWMOVLISLNS 722
DB 892 -----SPATETSDNI-SPVASPHITGSLVAF-----WVDARGGS 924
QY 723 QRRRLHSAASKLHLKSEGFMYKCEVELAMAMATIDSEFQONQRLVIYIDGLDACQ 782
DB 925 MRSRHN-----GLRAVIP-----PRTCAAPRTICRLV----- 953
QY 783 DKVLQMLDVRVLFSGKPIAIFASDPHIIKAINONLNSVLADSNINGDMWRN-IYHL 841
DB 954 --KPKLSTPPPL-----AEEGLASRIITALPGTGAQFISPIYVEI 992
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITCSPTGTQEDTDRVSONSLGEWTKLGSRTALN 900
DB 993 PHFASHGRG---DRELVLVLSSENGSV-----WHEHRSRYGESYLDQILN-GMDELG 1040
QY 901 RDTYRRRQOMRTTROMSFDLTKLVLT--EDMFSDISFQIMRLINITYVTGLLRAN 957
DB 1041 SLELEKRRVCRIIT--TDFPLFYVMSRLCOO-YDIIGPEG-----GSLKSKLVPLV 1090
QY 958 QIFENMDRLASWINTFQOMPYRTSWILYLEFTEGLDPDQMTLK-----TWYER 1005
DB 1091 QATFPENAVTKRKALQO-----AOPVDELVTKLGNQATFSPIVEPR 1136
QY 1006 ISK-----NIPTKDVEPLEIDGDIRNEVF-----LSSRTPLVA 1042
DB 1137 RRFKFRIGIRIPLPEWTQNPDSGSDFTSLRLCSYIGTDQAMWEDITGTLKYVA 1196
QY 1043 RDVKTPLPCVTNIDPKREIITADYRAREOINIGLAPPLPHGEPFPPSGYSOPASY 1102
DB 1197 NECANF---TTNV--SARFMLSOCPRFAEAVNATILYKEL-----TAVPYAKR 1241
QY 1103 CSSASFNGPPGCVVSPQPHSSYSSGLSGQHPPYNAANVPATGSSLLSMFVDPYCEK 1162
DB 1242 VIFKAKNDPRBG-----RLCYCMTDDKDYKT 1268
QY 1163 LRQ-----IEG-----LDONMPOYCTTIRKAN-----INGRVL 1191
DB 1269 LEQENFVVARSDIEVLGMSLFAELSGNLV---YKKAQOQSFHFQSPREKRLA 1323
QY 1192 SQCNIDELKE-----MANFGD-WHLFRSVMLEKMSVESQVPEPR----- 1233
DB 1324 MPVAVRDSSEPPGSLSLFKKAMKYEOTOHILCHLNTMPPCAGSAGERRRTPPLAL 1383
QY 1234 ---PLANENSSAPVPHGESARR-----SSHHELPLTESSQTPYTLNFSFELMTGLD-- 1283
DB 1384 RYSLIESTPGSLSTGTOAEKMAVISEHGLSMALARE---LOFVEDINIRVENP 1439
QY 1284 ---EGAPRHSNLMQOQTRTPSLSSLNODSSISEIKLTDQVA-----EYRD 1329
DB 1440 NSLLEQSVALLNT-WVIREGONAMNENLYTALQSIDGELYNMLEGGROSRLNKPRRH 1498
QY 1330 AYREYIAQMSQLEGGTSSITSGRSSPHSTYYIGOSSGGSINSTLEQERKEGELKOED 1389
DB 1499 TDRYSLSPSQMNG--YSLQDELISP-----ASIGCALSSPLRADQ----- 1538
QY 1390 GRKSFMLKRGVYIDYSSSGVSTNAPSLDPT---TEED---EKSDOS--GSKLLPKKS 1440
DB 1539 -----YWNEVAVYALDAIPLAATHTDWTLEMSDQVMSAGLTPSLVT 1578
QY 1441 SERPSIFQTDLTKGGLRYOKLPSDEDESGTGHVQJLTPHCSKMIFFKRLAKARECASP 1500
DB 1579 AEDSSL-----ECKAKEDSDMTGH-----EKKLBALSEPRGP 1612
QY 1501 QHSASAPITRTIKAKAYLSDA---LLDKKSSDSGVNSNESSPNHSLHNEAADDQLEKA 1557
DB 1613 ELGSLF---LVEDVDVSDATNGLIDLEL--QEGGORSSEKLPCKSRQDDATGAGQ--DSE 1666
QY 1558 NLELEDGSHGKGMHSHLSGLQDPIIARMSISSEDK-----KSPSCSLIASPPEEW 1612

```

```

DB 1667 NEVSLVSHQRQOANITHS-----PTVSQYTERSQDRLQWMDADGISTVYLDQAAQOSW 1720
RESULT 2
AC ANK2_HUMAN STANDARD: PRT: 3924 AA.
ID ANK2_HUMAN
AC 001484; 001485.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
RN [2]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordell E., Bennett V.;
RT "440-kD ankyrin: structure of the major developmentally regulated
domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
RN [4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Meininger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RL Lue S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
ankyrin gene.";
RL Genomics 10:858-866(1991).
CC - FUNCTION: Attach integral membrane proteins to cytoskeletal
elements. Also bind to cytoskeletal proteins.
CC - ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CELLS THROUGHOUT THE BRAIN.
CC - PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
AND FUNCTION (POTENTIAL).
CC - SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC - SIMILARITY: CONTAINS 23 ANK REPEATS.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X56957; CAA40278.1; -
CC EMBL; X56958; CAA40279.2; -
CC EMBL; Z26634; CAA42644.1; -
CC EMBL; M37123; AAA62828.1; -
CC PIR; S14533; S14533.
CC PIR; A39643; A39643.
CC PIR; B39643; B39643.
CC PIR; S14569; S14569.

```

DR HSP; P42771; IDC2.  
 DR Genew; HGNC:493; ANK2.  
 DR MIM; 106410; -.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000488; Death.  
 DR InterPro: IPR000906; Z05.  
 DR Pfam; PF00023; ank; 24.  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00791; Z05; 1.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 21.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00218; Z05; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 20.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 DR Cytochrome; Alternative splicing; Repeat; ANK repeat;  
 KW Phosphorylation.  
 FT REPEAT 63 92 ANK 1.  
 FT REPEAT 96 125 ANK 2.  
 FT REPEAT 129 158 ANK 3.  
 FT REPEAT 162 191 ANK 4.  
 FT REPEAT 193 220 ANK 5.  
 FT REPEAT 232 261 ANK 6.  
 FT REPEAT 265 294 ANK 7.  
 FT REPEAT 298 327 ANK 8.  
 FT REPEAT 331 360 ANK 9.  
 FT REPEAT 364 393 ANK 10.  
 FT REPEAT 397 426 ANK 11.  
 FT REPEAT 430 459 ANK 12.  
 FT REPEAT 463 492 ANK 13.  
 FT REPEAT 496 525 ANK 14.  
 FT REPEAT 529 558 ANK 15.  
 FT REPEAT 562 591 ANK 16.  
 FT REPEAT 595 624 ANK 17.  
 FT REPEAT 628 657 ANK 18.  
 FT REPEAT 661 690 ANK 19.  
 FT REPEAT 694 723 ANK 20.  
 FT REPEAT 727 756 ANK 21.  
 FT REPEAT 760 789 ANK 22.  
 FT REPEAT 793 822 ANK 23.  
 FT REPEAT 825 854 ANK 24.  
 FT REPEAT 857 886 ANK 25.  
 FT REPEAT 889 918 ANK 26.  
 FT REPEAT 921 950 ANK 27.  
 FT REPEAT 953 982 ANK 28.  
 FT REPEAT 985 1014 ANK 29.  
 FT REPEAT 1017 1046 ANK 30.  
 FT REPEAT 1049 1078 ANK 31.  
 FT REPEAT 1081 1110 ANK 32.  
 FT REPEAT 1113 1142 ANK 33.  
 FT REPEAT 1145 1174 ANK 34.  
 FT REPEAT 1177 1206 ANK 35.  
 FT REPEAT 1209 1238 ANK 36.  
 FT REPEAT 1241 1270 ANK 37.  
 FT REPEAT 1273 1302 ANK 38.  
 FT REPEAT 1305 1334 ANK 39.  
 FT REPEAT 1337 1366 ANK 40.  
 FT REPEAT 1369 1398 ANK 41.  
 FT REPEAT 1401 1430 ANK 42.  
 FT REPEAT 1433 1462 ANK 43.  
 FT REPEAT 1465 1494 ANK 44.  
 FT REPEAT 1497 1526 ANK 45.  
 FT REPEAT 1529 1558 ANK 46.  
 FT REPEAT 1561 1590 ANK 47.  
 FT REPEAT 1593 1622 ANK 48.  
 FT REPEAT 1625 1654 ANK 49.  
 FT REPEAT 1657 1686 ANK 50.  
 FT REPEAT 1689 1718 ANK 51.  
 FT REPEAT 1721 1750 ANK 52.  
 FT REPEAT 1753 1782 ANK 53.  
 FT REPEAT 1785 1814 ANK 54.  
 FT REPEAT 1817 1846 ANK 55.  
 FT REPEAT 1849 1878 ANK 56.  
 FT REPEAT 1881 1910 ANK 57.  
 FT REPEAT 1913 1942 ANK 58.  
 FT REPEAT 1945 1974 ANK 59.  
 FT REPEAT 1977 2006 ANK 60.  
 FT REPEAT 2009 2038 ANK 61.  
 FT REPEAT 2041 2070 ANK 62.  
 FT REPEAT 2073 2102 ANK 63.  
 FT REPEAT 2105 2134 ANK 64.  
 FT REPEAT 2137 2166 ANK 65.  
 FT REPEAT 2169 2198 ANK 66.  
 FT REPEAT 2201 2230 ANK 67.  
 FT REPEAT 2233 2262 ANK 68.  
 FT REPEAT 2265 2294 ANK 69.  
 FT REPEAT 2297 2326 ANK 70.  
 FT REPEAT 2329 2358 ANK 71.  
 FT REPEAT 2361 2390 ANK 72.  
 FT REPEAT 2393 2422 ANK 73.  
 FT REPEAT 2425 2454 ANK 74.  
 FT REPEAT 2457 2486 ANK 75.  
 FT REPEAT 2489 2518 ANK 76.  
 FT REPEAT 2521 2550 ANK 77.  
 FT REPEAT 2553 2582 ANK 78.  
 FT REPEAT 2585 2614 ANK 79.  
 FT REPEAT 2617 2646 ANK 80.  
 FT REPEAT 2649 2678 ANK 81.  
 FT REPEAT 2681 2710 ANK 82.  
 FT REPEAT 2713 2742 ANK 83.  
 FT REPEAT 2745 2774 ANK 84.  
 FT REPEAT 2777 2806 ANK 85.  
 FT REPEAT 2809 2838 ANK 86.  
 FT REPEAT 2841 2870 ANK 87.  
 FT REPEAT 2873 2902 ANK 88.  
 FT REPEAT 2905 2934 ANK 89.  
 FT REPEAT 2937 2966 ANK 90.  
 FT REPEAT 2969 2998 ANK 91.  
 FT REPEAT 3001 3030 ANK 92.  
 FT REPEAT 3033 3062 ANK 93.  
 FT REPEAT 3065 3094 ANK 94.  
 FT REPEAT 3097 3126 ANK 95.  
 FT REPEAT 3129 3158 ANK 96.  
 FT REPEAT 3161 3190 ANK 97.  
 FT REPEAT 3193 3222 ANK 98.  
 FT REPEAT 3225 3254 ANK 99.  
 FT REPEAT 3257 3286 ANK 100.  
 FT REPEAT 3289 3318 ANK 101.  
 FT REPEAT 3321 3350 ANK 102.  
 FT REPEAT 3353 3382 ANK 103.  
 FT REPEAT 3385 3414 ANK 104.  
 FT REPEAT 3417 3446 ANK 105.  
 FT REPEAT 3449 3478 ANK 106.  
 FT REPEAT 3481 3510 ANK 107.  
 FT REPEAT 3513 3542 ANK 108.  
 FT REPEAT 3545 3574 ANK 109.  
 FT REPEAT 3577 3606 ANK 110.  
 FT REPEAT 3609 3638 ANK 111.  
 FT REPEAT 3641 3670 ANK 112.  
 FT REPEAT 3673 3702 ANK 113.  
 FT REPEAT 3705 3734 ANK 114.  
 FT REPEAT 3737 3766 ANK 115.  
 FT REPEAT 3769 3798 ANK 116.  
 FT REPEAT 3801 3830 ANK 117.  
 FT REPEAT 3833 3862 ANK 118.  
 FT REPEAT 3865 3894 ANK 119.  
 FT REPEAT 3897 3926 ANK 120.  
 FT REPEAT 3929 3958 ANK 121.  
 FT REPEAT 3961 3990 ANK 122.  
 FT REPEAT 3993 4022 ANK 123.  
 FT REPEAT 4025 4054 ANK 124.  
 FT REPEAT 4057 4086 ANK 125.  
 FT REPEAT 4089 4118 ANK 126.  
 FT REPEAT 4121 4150 ANK 127.  
 FT REPEAT 4153 4182 ANK 128.  
 FT REPEAT 4185 4214 ANK 129.  
 FT REPEAT 4217 4246 ANK 130.  
 FT REPEAT 4249 4278 ANK 131.  
 FT REPEAT 4281 4310 ANK 132.  
 FT REPEAT 4313 4342 ANK 133.  
 FT REPEAT 4345 4374 ANK 134.  
 FT REPEAT 4377 4406 ANK 135.  
 FT REPEAT 4409 4438 ANK 136.  
 FT REPEAT 4441 4470 ANK 137.  
 FT REPEAT 4473 4502 ANK 138.  
 FT REPEAT 4505 4534 ANK 139.  
 FT REPEAT 4537 4566 ANK 140.  
 FT REPEAT 4569 4598 ANK 141.  
 FT REPEAT 4601 4630 ANK 142.  
 FT REPEAT 4633 4662 ANK 143.  
 FT REPEAT 4665 4694 ANK 144.  
 FT REPEAT 4697 4726 ANK 145.  
 FT REPEAT 4729 4758 ANK 146.  
 FT REPEAT 4761 4790 ANK 147.  
 FT REPEAT 4793 4822 ANK 148.  
 FT REPEAT 4825 4854 ANK 149.  
 FT REPEAT 4857 4886 ANK 150.  
 FT REPEAT 4889 4918 ANK 151.  
 FT REPEAT 4921 4950 ANK 152.  
 FT REPEAT 4953 4982 ANK 153.  
 FT REPEAT 4985 5014 ANK 154.  
 FT REPEAT 5017 5046 ANK 155.  
 FT REPEAT 5049 5078 ANK 156.  
 FT REPEAT 5081 5110 ANK 157.  
 FT REPEAT 5113 5142 ANK 158.  
 FT REPEAT 5145 5174 ANK 159.  
 FT REPEAT 5177 5206 ANK 160.  
 FT REPEAT 5209 5238 ANK 161.  
 FT REPEAT 5241 5270 ANK 162.  
 FT REPEAT 5273 5302 ANK 163.  
 FT REPEAT 5305 5334 ANK 164.  
 FT REPEAT 5337 5366 ANK 165.  
 FT REPEAT 5369 5398 ANK 166.  
 FT REPEAT 5401 5430 ANK 167.  
 FT REPEAT 5433 5462 ANK 168.  
 FT REPEAT 5465 5494 ANK 169.  
 FT REPEAT 5497 5526 ANK 170.  
 FT REPEAT 5529 5558 ANK 171.  
 FT REPEAT 5561 5590 ANK 172.  
 FT REPEAT 5593 5622 ANK 173.  
 FT REPEAT 5625 5654 ANK 174.  
 FT REPEAT 5657 5686 ANK 175.  
 FT REPEAT 5689 5718 ANK 176.  
 FT REPEAT 5721 5750 ANK 177.  
 FT REPEAT 5753 5782 ANK 178.  
 FT REPEAT 5785 5814 ANK 179.  
 FT REPEAT 5817 5846 ANK 180.  
 FT REPEAT 5849 5878 ANK 181.  
 FT REPEAT 5881 5910 ANK 182.  
 FT REPEAT 5913 5942 ANK 183.  
 FT REPEAT 5945 5974 ANK 184.  
 FT REPEAT 5977 6006 ANK 185.  
 FT REPEAT 6009 6038 ANK 186.  
 FT REPEAT 6041 6070 ANK 187.  
 FT REPEAT 6073 6102 ANK 188.  
 FT REPEAT 6105 6134 ANK 189.  
 FT REPEAT 6137 6166 ANK 190.  
 FT REPEAT 6169 6198 ANK 191.  
 FT REPEAT 6201 6230 ANK 192.  
 FT REPEAT 6233 6262 ANK 193.  
 FT REPEAT 6265 6294 ANK 194.  
 FT REPEAT 6297 6326 ANK 195.  
 FT REPEAT 6329 6358 ANK 196.  
 FT REPEAT 6361 6390 ANK 197.  
 FT REPEAT 6393 6422 ANK 198.  
 FT REPEAT 6425 6454 ANK 199.  
 FT REPEAT 6457 6486 ANK 200.  
 FT REPEAT 6489 6518 ANK 201.  
 FT REPEAT 6521 6550 ANK 202.  
 FT REPEAT 6553 6582 ANK 203.  
 FT REPEAT 6585 6614 ANK 204.  
 FT REPEAT 6617 6646 ANK 205.  
 FT REPEAT 6649 6678 ANK 206.  
 FT REPEAT 6681 6710 ANK 207.  
 FT REPEAT 6713 6742 ANK 208.  
 FT REPEAT 6745 6774 ANK 209.  
 FT REPEAT 6777 6806 ANK 210.  
 FT REPEAT 6809 6838 ANK 211.  
 FT REPEAT 6841 6870 ANK 212.  
 FT REPEAT 6873 6902 ANK 213.  
 FT REPEAT 6905 6934 ANK 214.  
 FT REPEAT 6937 6966 ANK 215.  
 FT REPEAT 6969 6998 ANK 216.  
 FT REPEAT 7001 7030 ANK 217.  
 FT REPEAT 7033 7062 ANK 218.  
 FT REPEAT 7065 7094 ANK 219.  
 FT REPEAT 7097 7126 ANK 220.  
 FT REPEAT 7129 7158 ANK 221.  
 FT REPEAT 7161 7190 ANK 222.  
 FT REPEAT 7193 7222 ANK 223.  
 FT REPEAT 7225 7254 ANK 224.  
 FT REPEAT 7257 7286 ANK 225.  
 FT REPEAT 7289 7318 ANK 226.  
 FT REPEAT 7321 7350 ANK 227.  
 FT REPEAT 7353 7382 ANK 228.  
 FT REPEAT 7385 7414 ANK 229.  
 FT REPEAT 7417 7446 ANK 230.  
 FT REPEAT 7449 7478 ANK 231.  
 FT REPEAT 7481 7510 ANK 232.  
 FT REPEAT 7513 7542 ANK 233.  
 FT REPEAT 7545 7574 ANK 234.  
 FT REPEAT 7577 7606 ANK 235.  
 FT REPEAT 7609 7638 ANK 236.  
 FT REPEAT 7641 7670 ANK 237.  
 FT REPEAT 7673 7702 ANK 238.  
 FT REPEAT 7705 7734 ANK 239.  
 FT REPEAT 7737 7766 ANK 240.  
 FT REPEAT 7769 7798 ANK 241.  
 FT REPEAT 7801 7830 ANK 242.  
 FT REPEAT 7833 7862 ANK 243.  
 FT REPEAT 7865 7894 ANK 244.  
 FT REPEAT 7897 7926 ANK 245.  
 FT REPEAT 7929 7958 ANK 246.  
 FT REPEAT 7961 7990 ANK 247.  
 FT REPEAT 7993 8022 ANK 248.  
 FT REPEAT 8025 8054 ANK 249.  
 FT REPEAT 8057 8086 ANK 250.  
 FT REPEAT 8089 8118 ANK 251.  
 FT REPEAT 8121 8150 ANK 252.  
 FT REPEAT 8153 8182 ANK 253.  
 FT REPEAT 8185 8214 ANK 254.  
 FT REPEAT 8217 8246 ANK 255.  
 FT REPEAT 8249 8278 ANK 256.  
 FT REPEAT 8281 8310 ANK 257.  
 FT REPEAT 8313 8342 ANK 258.  
 FT REPEAT 8345 8374 ANK 259.  
 FT REPEAT 8377 8406 ANK 260.  
 FT REPEAT 8409 8438 ANK 261.  
 FT REPEAT 8441 8470 ANK 262.  
 FT REPEAT 8473 8502 ANK 263.  
 FT REPEAT 8505 8534 ANK 264.  
 FT REPEAT 8537 8566 ANK 265.  
 FT REPEAT 8569 8598 ANK 266.  
 FT REPEAT 8601 8630 ANK 267.  
 FT REPEAT 8633 8662 ANK 268.  
 FT REPEAT 8665 8694 ANK 269.  
 FT REPEAT 8697 8726 ANK 270.  
 FT REPEAT 8729 8758 ANK 271.  
 FT REPEAT 8761 8790 ANK 272.  
 FT REPEAT 8793 8822 ANK 273.  
 FT REPEAT 8825 8854 ANK 274.  
 FT REPEAT 8857 8886 ANK 275.  
 FT REPEAT 8889 8918 ANK 276.  
 FT REPEAT 8921 8950 ANK 277.  
 FT REPEAT 8953 8982 ANK 278.  
 FT REPEAT 8985 9014 ANK 279.  
 FT REPEAT 9017 9046 ANK 280.  
 FT REPEAT 9049 9078 ANK 281.  
 FT REPEAT 9081 9110 ANK 282.  
 FT REPEAT 9113 9142 ANK 283.  
 FT REPEAT 9145 9174 ANK 284.  
 FT REPEAT 9177 9206 ANK 285.  
 FT REPEAT 9209 9238 ANK 286.  
 FT REPEAT 9241 9270 ANK 287.  
 FT REPEAT 9273 9302 ANK 288.  
 FT REPEAT 9305 9334 ANK 289.  
 FT REPEAT 9337 9366 ANK 290.  
 FT REPEAT 9369 9398 ANK 291.  
 FT REPEAT 9401 9430 ANK 292.  
 FT REPEAT 9433 9462 ANK 293.  
 FT REPEAT 9465 9494 ANK 294.  
 FT REPEAT 9497 9526 ANK 295.  
 FT REPEAT 9529 9558 ANK 296.  
 FT REPEAT 9561 9590 ANK 297.  
 FT REPEAT 9593 9622 ANK 298.  
 FT REPEAT 9625 9654 ANK 299.  
 FT REPEAT 9657 9686 ANK 300.  
 FT REPEAT 9689 9718 ANK 301.  
 FT REPEAT 9721 9750 ANK 302.  
 FT REPEAT 9753 9782 ANK 303.  
 FT REPEAT 9785 9814 ANK 304.  
 FT REPEAT 9817 9846 ANK 305.  
 FT REPEAT 9849 9878 ANK 306.  
 FT REPEAT 9881 9910 ANK 307.  
 FT REPEAT 9913 9942 ANK 308.  
 FT REPEAT 9945 9974 ANK 309.  
 FT REPEAT 9977 10006 ANK 310.  
 FT REPEAT 10009 10038 ANK 311.  
 FT REPEAT 10041 10070 ANK 312.  
 FT REPEAT 10073 10102 ANK 313.  
 FT REPEAT 10105 10134 ANK 314.  
 FT REPEAT 10137 10166 ANK 315.  
 FT REPEAT 10169 10198 ANK 316.  
 FT REPEAT 10201 10230 ANK 317.  
 FT REPEAT 10233 10262 ANK 318.  
 FT REPEAT 10265 10294 ANK 319.  
 FT REPEAT 10297 10326 ANK 320.  
 FT REPEAT 10329 10358 ANK 321.  
 FT REPEAT 10361 10390 ANK 322.  
 FT REPEAT 10393 10422 ANK 323.  
 FT REPEAT 10425 10454 ANK 324.  
 FT REPEAT 10457 10486 ANK 325.  
 FT REPEAT 10489 10518 ANK 326.  
 FT REPEAT 10521 10550 ANK 327.  
 FT REPEAT 10553 10582 ANK 328.  
 FT REPEAT 10585 10614 ANK 329.  
 FT REPEAT 10617 10646 ANK 330.  
 FT REPEAT 10649 10678 ANK 331.  
 FT REPEAT 10681 10710 ANK 332.  
 FT REPEAT 10713 10742 ANK 333.  
 FT REPEAT 10745 10774 ANK 334.  
 FT REPEAT 10777 10806 ANK 335.  
 FT REPEAT 10809 10838 ANK 336.  
 FT REPEAT 10841 10870 ANK 337.  
 FT REPEAT 10873 10902 ANK 338.  
 FT REPEAT 10905 10934 ANK 339.  
 FT REPEAT 10937 10966 ANK 340.  
 FT REPEAT 10969 10998 ANK 341.  
 FT REPEAT 11001 11030 ANK 342.  
 FT REPEAT 11033 11062 ANK 343.  
 FT REPEAT 11065 11094 ANK 344.  
 FT REPEAT 11097 11126 ANK 345.  
 FT REPEAT 11129 11158 ANK 346.  
 FT REPEAT 11161 11190 ANK 347.  
 FT REPEAT 11193 11222 ANK 348.  
 FT REPEAT 11225 11254 ANK 349.  
 FT REPEAT 11257 11286 ANK 350.  
 FT REPEAT 11289 11318 ANK 351.  
 FT REPEAT 11321 11350 ANK 352.  
 FT REPEAT 11353 11382 ANK 353.  
 FT REPEAT 11385 11414 ANK 354.  
 FT REPEAT 11417 11446 ANK 355.  
 FT REPEAT 11449 11478 ANK 356.  
 FT REPEAT 11481 11510 ANK 357.  
 FT REPEAT 11513 11542 ANK 358.  
 FT REPEAT 11545 11574 ANK 359.  
 FT REPEAT 11577 11606 ANK 360.  
 FT REPEAT 11609 11638 ANK 361.  
 FT REPEAT 11641 11670 ANK 362.  
 FT REPEAT 11673 11702 ANK 363.  
 FT REPEAT 11705 11734 ANK 364.  
 FT REPEAT 11737 11766 ANK 365.  
 FT REPEAT 11769 11798 ANK 366.  
 FT REPEAT 11801 11830 ANK 367.  
 FT REPEAT 11833 11862 ANK 368.  
 FT REPEAT 11865 11894 ANK 369.  
 FT REPEAT 11897 11926 ANK 370.  
 FT REPEAT 11929 11958 ANK 371.  
 FT REPEAT 11961 11990 ANK 372.  
 FT REPEAT 11993 12022 ANK 373.  
 FT REPEAT 12025 12054 ANK 374.  
 FT REPEAT 12057 12086 ANK 375.  
 FT REPEAT 12089 12118 ANK 376.  
 FT REPEAT 12121 12150 ANK 377.  
 FT REPEAT 12153 12182 ANK 378.  
 FT REPEAT 12185 12214 ANK 379.  
 FT REPEAT 12217 12246 ANK 380.  
 FT REPEAT 12249 12278 ANK 381.  
 FT REPEAT 12281 12310 ANK 382.  
 FT REPEAT 12313 12342 ANK 383.  
 FT REPEAT 12345 12374 ANK 384.  
 FT REPEAT 12377 12406 ANK 385.  
 FT REPEAT 12409 12438 ANK 386.  
 FT REPEAT 12441 12470 ANK 387.  
 FT REPEAT 12473 12502 ANK 388.  
 FT REPEAT 12505 12534 ANK 389.  
 FT REPEAT 12537 12566 ANK 390.  
 FT REPEAT 12569 12598 ANK 391.  
 FT REPEAT 12601 12630 ANK 392.  
 FT REPEAT 12633 12662 ANK 393.  
 FT REPEAT 12665 12694 ANK 394.  
 FT REPEAT 12697 12726 ANK 395.  
 FT REPEAT 12729 12758 ANK 396.  
 FT REPEAT 12761 12790 ANK 397.  
 FT REPEAT 12793 12822 ANK 398.  
 FT REPEAT 12825 12854 ANK 399.  
 FT REPEAT 12857 12886 ANK 400.  
 FT REPEAT 12889 12918 ANK 401.  
 FT REPEAT 12921 12950 ANK 402.  
 FT REPEAT 12953 12982 ANK 403.  
 FT REPEAT 12985 13014 ANK 404.  
 FT REPEAT 13017 13046 ANK 405.  
 FT REPEAT 13049 13078 ANK 406.  
 FT REPEAT 13081 13110 ANK 407.  
 FT REPEAT 13113 13142 ANK 408.  
 FT REPEAT 13145 13174 ANK 409.  
 FT REPEAT 13177 13206 ANK 410.  
 FT REPEAT 13209 13238 ANK 411.  
 FT REPEAT 13241 13270 ANK 412.  
 FT REPEAT 13273 13302 ANK 413.  
 FT REPEAT 13305 13334 ANK 414.  
 FT REPEAT 13337 13366 ANK 415.  
 FT REPEAT 13369 13398 ANK 416.  
 FT REPEAT 13401 13430 ANK 417.  
 FT REPEAT 13433 13462 ANK 418.  
 FT REPEAT 13465 13494 ANK 419.  
 FT REPEAT 13497 13526 ANK 420.  
 FT REPEAT 13529 13558 ANK 421.  
 FT REPEAT 13561 13590 ANK 422.  
 FT REPEAT 13593 13622 ANK 423.  
 FT REPEAT 13625 13654 ANK 424.  
 FT REPEAT 13657 13686 ANK 425.  
 FT REPEAT 13689 13718 ANK 426.  
 FT REPEAT 13721 13750 ANK 427.  
 FT REPEAT 13753 13782 ANK 428.  
 FT REPEAT 13785 13814 ANK 429.  
 FT REPEAT 13817 13846 ANK 430.  
 FT REPEAT 13849 13878 ANK 431.  
 FT REPEAT 13881 13910 ANK 432.  
 FT REPEAT 13913 13942 ANK 433.  
 FT REPEAT 13945 13974 ANK 434.  
 FT REPEAT 13977 14006 ANK 435.  
 FT REPEAT 14009 14038 ANK 436.  
 FT REPEAT 14041 14070 ANK 437.  
 FT REPEAT 14073 14102 ANK 438.  
 FT REPEAT 14105 14134 ANK 439.  
 FT REPEAT 14137 14166 ANK 440.  
 FT REPEAT 14169 14198 ANK 441.  
 FT REPEAT 14201 14230 ANK 442.  
 FT REPEAT 14233 14262 ANK 443.  
 FT REPEAT 14265 14294 ANK 444.  
 FT REPEAT 14297 14326 ANK 445.  
 FT REPEAT 14329 14358 ANK 446.  
 FT REPEAT 14361 14390 ANK 447.  
 FT REPEAT 14393 14422 ANK 448.  
 FT REPEAT 14425 14454 ANK 449.  
 FT REPEAT 14457 14486 ANK 450.  
 FT REPEAT 14489 14518 ANK 451.  
 FT REPEAT 14521 14550 ANK 452.  
 FT REPEAT 14553 14582 ANK 453.  
 FT REPEAT 14585 14614 ANK 454.  
 FT REPEAT 14617 14646 ANK 455.  
 FT REPEAT 14649 14678 ANK 456.  
 FT REPEAT 14681 14710 ANK 457.  
 FT REPEAT 14713 14742 ANK 458.  
 FT REPEAT 14745 14774 ANK 459.  
 FT REPEAT 14777 14806 ANK 460.  
 FT REPEAT 14809 14838 ANK 461.  
 FT REPEAT 14841 14870 ANK 462.  
 FT REPEAT 14873 14902 ANK 463.  
 FT REPEAT 14905 14934 ANK 464.  
 FT REPEAT 14937 14966 ANK 465.  
 FT REPEAT 14969 14998 ANK 466.  
 FT REPEAT 15001 15030 ANK 467.  
 FT REPEAT 15033 15062 ANK 468.  
 FT REPEAT 15065 15094 ANK 469.  
 FT REPEAT 15097 15126 ANK 470.  
 FT REPEAT 15129 15158 ANK 471.  
 FT REPEAT 15161 15190 ANK 472.  
 FT REPEAT 15193 15222 ANK 473.  
 FT REPEAT 15225 15254 ANK 474.  
 FT REPEAT 15257 15286 ANK 475.  
 FT REPEAT 15289 15318 ANK 476.  
 FT REPEAT 15321 15350 ANK 477.  
 FT REPEAT 15353 15382 ANK 478.  
 FT REPEAT 15385 15414 ANK 479.  
 FT REPEAT 15417 15446 ANK 480.  
 FT REPEAT 15449 15478 ANK 481.  
 FT REPEAT 15481 15510 ANK 482.  
 FT REPEAT 15513 15542 ANK 483.  
 FT REPEAT 15545 15574 ANK 484.  
 FT REPEAT 15577 15606 ANK 485.  
 FT REPEAT 15609 15638 ANK 486.  
 FT REPEAT 15641 15670 ANK



FT REPEAT 465 494 ANK 14.  
 FT REPEAT 498 527 ANK 15.  
 FT REPEAT 531 560 ANK 16.  
 FT REPEAT 564 593 ANK 17.  
 FT REPEAT 597 626 ANK 18.  
 FT REPEAT 630 659 ANK 19.  
 FT REPEAT 663 692 ANK 20.  
 FT REPEAT 696 725 ANK 21.  
 FT REPEAT 729 758 ANK 22.  
 FT REPEAT 762 791 ANK 23.  
 FT DOMAIN 1399 1483 DEATH.  
 SQ SEQUENCE 1862 AA; 204242 MW; AE6B85B5B29001E5 CRC64;

Query Match 6.0%; Score 527.5; DB 1; Length 1862;  
 Best Local Similarity 20.0%; Pred. No. 4.5e-21;  
 Matches 407; Conservative 279; Mismatches 661; Indels 683; Gaps 83;

QY 21 LKALLECKDNERGCGOTPLMLAEONGVEIYKELKNGANCNLEDJMDMTALISASK 80  
 DB 90 VRELNYGAMVNAOQSGFTPLYMAOENHLEVVYKLENGANQVATEDEGFTPLVALQ 149  
 QY 81 EGH  
 DB 150 QCHENWVAHLNYGKGVRLPALHIAARNDTRTAVALQNDPNDVLSKTGFTPLHIA 209  
 QY 84 -----IHVEELKSGASLEHRDMGWTALMAGYKGRDVEVLLSHGAMPSTGLQYS 138  
 DB 210 AHXENLVNAQOLLNNGASVNTFPONGITPLHIASRGVIMVRLLDGKAQIE-TRKDE 268  
 QY 139 YPTITMAAGRGHADIYHLLONGAKVNSDKYGTTPVMAARKGHECVKHLAMGADV 198  
 DB 269 LTPHCAARNGHVRISLIDHGAPIOAKTKNGLSHMAOQHDLCVRLLOYNAEID 328  
 QY 199 -----ORGA-----NSMTALYAVKGGYTSQVKEIK 225  
 DB 329 DITDLHLPLHYAACHHRAKAYLLDCKAPNSRHALGFTPLHACKKNIRVWELLK 388  
 QY 226 RNPVNLTDKGNATLMIASKEGHEIYODLLDAGTYNIPDRSGDVLIGAVRGHVEI 285  
 DB 389 TGASIDAVTESGLFPLHYASFHGLPLVKNLLQNGASPNVSNVPEPLHMAAAGHTEV 448  
 QY 286 VRALLOKADIDIRQDKKALHYAIVEKGNATWYRDIIQCPDIEICTKQDEFTPLIKATK 345  
 DB 449 AKYLLONKAKANAKAKDQTPHCAARIGHTGMVKLENGASPNLATTACHTDLHNAAR 508  
 QY 346 MRNIEVELLDKGAKYSAVDKGDTPHVAIRGRSRRLAELLRNPDGRLLTRPKAG 405  
 DB 509 EGHVDTALALKEKASQACMTKKGFTPLHVAKAGYKRLABELLEHDA-----HPNAG 562  
 QY 406 E---TPYIDCSHO-----KSIL-----TOIFGAR----- 427  
 DB 563 KNGLPLPHVAHHNNLDIVKLLIPRGSGSPHSANNGYTPLHIAKQOIEVARSLQYGG 622  
 QY 428 -----HLSPETDGMGYDLYSSA-----LADI 451  
 DB 623 SANASVOGVTPLHIAOEGHTEWALLLSKQANGNLGNKSGTLPLHLSVOEGHVALDV 682  
 QY 452 LSE-PTMOPICV-----LYAOWSGK-SPLIKLIED-EMKTFAGOOTEPPLFQPSWL 501  
 DB 683 LIKHGVYVDATTRMGYPLPHVASHYGNIKLVKFLLOQADVNAKTKIGY-SPLHQAAQ 740  
 QY 502 ---IYFLLILLCGGIGLGFAPFVD-----TNLAIAISLPLALYIFFIYIYRGREG 553  
 DB 741 GHTDIVYLLKNG-----ASPNEVSSNGTTPLAIAKLGYSYDVLYKVT----- 786  
 QY 554 SMNMAALSTRLARHIGYLELLFLMFVNPELPEQF-TKALPYRFLFTYINRLSVGE 612  
 DB 787 ---DTSVYVSDKRMASIPETVDEILDYSEDEGLVSKA-----BRDSRDGEE 836  
 QY 613 TSAEMTATISDACEREGFLATRLR-----VFTE-ESQKKKKKKTCLPSFV 662  
 DB 837 KELDFVFKLDQVVE-----SPAIPRIPCVTPETVIRSEDDQAKKEVDEDSLIPS-- 888

QY 663 IFLEFYCITAGITLAIIFRPDKHLTVNAIILISVGLAFVLCRTFMVQVDSLNS 722  
 DB 889 -----SPATERSDNI-SPVASPVHTGLVSEF-----MVDARGGS 921  
 QY 723 ORKRLHSAKHLKSKSEGFMYLKECEVELMAMNATITSEFQNOFRLLVYIIDGACAG 782  
 DB 922 MGRSRH-----GLRVYIP-----PTCAPIRITRRLV----- 950  
 QY 783 DKYLOMLDITVRVLESGKPFIAFASDPHIIKAINONLSVLDNSINGHDYRN-IYH 841  
 DB 951 -KROKLTNPPL-----AEEGLASRIITALPTGAQFSPYIVEI 989  
 QY 842 PVFL-NRGLSNARKFLYTSATNGDITCSPTTQODTDRVSONSLGEMTLGSTALN 900  
 DB 990 PHRASRG---DRELVLRSNGSV-----WKEHRSRGESTYLDQILN-GMDEEG 1037  
 QY 901 RRDYRRRQMTITRQMSFDLKLVT-----EDWFSIDISPOWRRLNIVSYGRLLRAN 957  
 DB 1038 SLELEKRRVCRLIT--TDPELYFVIMSRLOQ-YDTIGPEG-----GSLRSKLVPLV 1087  
 QY 958 QITFNDRLASWNLTEQWYRTSWLILEETEGLPDQMTLKTMYERISKNIPTTKVYE 1017  
 DB 1088 QATPEPNAVTKKVALQ-----AQPVDELVTYKLLGNO----- 1121  
 QY 1018 PLEIDQDINFVEPLSSRTPVLVARDVTFPLCTVNLDPKREITADYRAAREQINIG 1077  
 DB 1122 -----ATFSP-IYTVERRKKRHPF-----G 1142  
 QY 1078 LAYPLPLEHGGPPRPSGYSPAS---VSSASFNCPFGGVSPQPHSSYSGLSGPOH 1134  
 DB 1143 LRIPLPSPWTDNR-DSGEDTSLRLCSVI-----GMDQAO-----MEDITGTJK 1189  
 QY 1135 PFY-NRAA-----VPATG-----SILLSMTYDVYCEKRLQIEGLDQNM 1174  
 DB 1190 LIYANECANFTTVASARFMLSDCPTAEVHRTATLLYKELTA-----V 1232  
 QY 1175 POYCTTIKKNANIN---GVLISQCNIDE-LKREMANFEDMHLFRSMVLEM-----RS 1222  
 DB 1233 PYAKVYVITAKMNDAREGRRCYCMTDVDVDTLQHEHFEVARSROIEVEGMPLEAF 1292  
 QY 1223 VESQVVP-----EDRFLNENSAAPVPHGESARR-----SSH-- 1254  
 DB 1293 LSGNLVYVKKAAQORSEHFQSPFENRLAIPVYVROSSRBPGLSFLRKTMYEDTOHIL 1352  
 QY 1255 -----TELPTELES-----SOTPYLNFSEELTDLGDAVR-----HSL 1292  
 DB 1353 CHNITMPCTKSGAEDRRRTITPLTKYSLISESRLEFTSDTDRVRRMAVIREHGL 1412  
 QY 1293 SWQSGTR-----RTPSLSLNSQDSSI-----ETSKLTKVQAEYRDAY 1331  
 DB 1413 SWAELARELOFVEDINRIREVPNSLDQSTALLTLWVDRGENAKM-ENLYTALRND 1471  
 QY 1332 REYIAQMSQEGTSGTSTISGRSSPHSYTYIQOSSSGGISHTLDOERKEBELQOEDGR 1391  
 DB 1472 RSEIVNM-LEG-----SGROS-----RNLKPER-----RHGD 1497  
 QY 1392 KSELMKRGVIVYSS-----SGVSTNEASPLDPTIEDEKSDQSGSKLPKKSSEPRS 1445  
 DB 1498 REYSLSPOVNGYSSLODELLSPASIQVALLP-SPLCAQYMEVYVIDAIPLAATEHDTM 1556  
 QY 1446 LFQTDILKLGGLRQKPLRDESGTGRVQITPHCSKMITRKLAKARECASPOEHA 1505  
 DB 1557 LEWSDQVMSAGILT-PSLVTABDS-----LECSKAEDSDA 1591  
 QY 1506 EPIRPTIKAKEXYLLDLDKSSQGVANSNESSPHSLHNAAD-DSOLEKANILIED 1564  
 DB 1592 IP-----EKLEGAHSEDT--QGPGLSQDLYVEDTVDSDATNGLADLAGEE 1637  
 QY 1565 EGHSGKGMPHSLGLOPFIIRMSICSEDK-----SPSECSLI----- 1604  
 DB 1638 GORSEKK--RQVSGEGDQTEFEVLSGQGVVIAHRTIDSPSVYRVDRSQARLDMQKQ 1695  
 QY 1605 ---ASSPEE---SW--PACQKAYNLNRTPSVTVLNNVTAPLRANONFDEI----- 1647

DB 1696 GSTAVHPQATSSWQEEVTOGPHSFORITTT-----OGPPGALQOETEVOVLVSTREHV 1750  
QY 1648 -EGIRETSOVLIRPGSPNPRTA-----YONENLKSMAHRKRSORSST 1688  
DB 1751 QNGPPTGSP--KAGKEPSLMAPESAFSGEOGDELQNGQVVEEQT 1798

RESULT 4  
ID ANK3\_HUMAN STANDARD; PRT; 4377 AA.  
AC 012955;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ankyrin 3 (ANK-3) (Ankyrin G).  
GN ANK3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain stem;  
RX MEDLINE=95138209; PubMed=7836469;  
RA Kordeli E., Lambert S., Bennett V.  
RT "Ankyrin. A new ankyrin gene with neural-specific isoforms localized  
RT at the axonal initial segment and node of Ranvier".  
RL J. Biol. Chem. 270:2352-2359(1995).  
CC -1- FUNCTION: Membrane-cytoskeleton linker.  
CC -1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by  
CC alternative splicing.  
CC -1- TISSUE SPECIFICITY: Expressed in brain and other tissues.  
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.  
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: U13616; AAA64834.1; -  
DR HSP: P53273; 1B18.  
DR Genew: HGNC:494; ANK3.  
DR MIM: 600465; -  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000488; Death.  
DR InterPro: IPR009066; ZUS.  
DR Pfam: PF00023; ank; 24.  
DR Pfam: PF00531; death; 1.  
DR Pfam: PF00791; ZUS; 1.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00248; ANK; 21.  
DR SMART: SM00005; DEATH; 1.  
DR SMART: SM00218; ZUS; 1.  
DR PROSITE: PS50088; ANK\_REPEAT; 21.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE: PS50017; DEATH\_DOMAIN; 1.  
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.

FT REPEAT 73 102 ANK 1.  
FT REPEAT 106 135 ANK 2.  
FT REPEAT 139 168 ANK 3.  
FT REPEAT 172 201 ANK 4.  
FT REPEAT 203 230 ANK 5.  
FT REPEAT 234 263 ANK 6.  
FT REPEAT 267 296 ANK 7.  
FT REPEAT 300 329 ANK 8.  
FT REPEAT 333 362 ANK 9.  
FT REPEAT 366 395 ANK 10.  
FT REPEAT 399 428 ANK 11.

FT REPEAT 432 461 ANK 12.  
FT REPEAT 465 494 ANK 13.  
FT REPEAT 498 527 ANK 14.  
FT REPEAT 531 560 ANK 15.  
FT REPEAT 564 593 ANK 16.  
FT REPEAT 597 626 ANK 17.  
FT REPEAT 630 659 ANK 18.  
FT REPEAT 663 692 ANK 19.  
FT REPEAT 696 725 ANK 20.  
FT REPEAT 729 758 ANK 21.  
FT REPEAT 762 791 ANK 22.  
FT REPEAT 795 825 ANK 23.  
FT DOMAIN 1519 1898 SER-RICH.  
FT DOMAIN 4090 4174 DEATH.  
SQ SEQUENCE 4377 AA; 480399 MW; F42379E55768B684 CRC64;

Query Match 5.9%; Score 526; DB 1; Length 4377;  
Best Local Similarity 20.0%; Pred. No. 2,1e-20;  
Matches 416; Conservative 266; Mismatches 664; Indels 734; Gaps 80;

QY 14 EEENIPALKALLEKCKVDENEGQPTPLMLAEGNVEIYKELKNGANCNLEDLQNT 73  
DB 83 KGEHVEVSELLQREANVDAATKGNLTALHTASLAGAEVVKVLTNGANVNAOSQNGFT 142  
QY 74 ALISAKEGHIIHVELEKSGASLEHRDMSGWTALMAVCYKRTDVELL----- 124  
DB 143 PLVMAAENHLEVFLLDNGASQSLATEDGFTPLVAVLQGHQVYSLLENDTKGVR 202  
QY 125 -----SHGANPSYTG----- 135  
DB 203 LPAHTAARKDDTKAAALLLQDNNAVDESKGFTPLHIAHYGINVATLLNRAAVD 262  
QY 136 ---QSYVPIIWAARGHADIYHLLQNGAVNCSKDTGTPVWAAKRG----- 182  
DB 263 FPARNDITPLHVASRGNANVKKLLDNGAKIDATRGTLPLHGASRGHEQVEMLLD 322  
QY 183 -----HLECYHLLAMGADVDOGANSMTALIVAVRGYTQS 219  
DB 323 RAAPILSKTKNGLSPLHMAQTGDHLCYQVLLQHNVPYDDVTDYTLVHAAGCHYKV 382  
QY 220 VEEILKRNPNVNLTKDGTALMISKE----- 247  
DB 383 AVLLDKKRNPAKALNGFTPLHIAKKNRILKVMELLKHGASIQAVTESGLTPHVAF 442  
QY 248 -GHIEIYODLDAGTYVNPDRSGDPLVIGAVRGVEIVRALLQKYADIDIRGDNKTA 306  
DB 443 KGHVNIQSOLHHRGASPTNTNVRGETALHMAARSGAEVVRIVYDGDAGVEKAKDDQTP 502  
QY 307 LYWAVEKGNATVVRDILQ--CNPD-----ELCT 333  
DB 503 LHTSARLGKADIVQQLQOGASPNMATSGYTPPLHSAREGHEVDAEFLDHGASLSITT 562  
QY 334 KQGEPLIKATKMRNIEVEVELLDGAKYSANDKGDPLHVAINGSRRLAELLRMPK 393  
DB 563 KKGFTPLHVAARYKLEVANLLQKSASPDAAKSGTLPLHVAHYDQKVALLL--D 619  
QY 394 DGRILYRPKAGETPNIDC-SHOKSILTOI--FGAR-----HLSPTEDGD 438  
DB 620 QGASPLAAKKNQYTPPLHIAKKNQMDIATLTLEYGDANAVYTRQGIASVHLAAGSHDM 679  
QY 439 LGDYLYSSALDILSEPTMOP-----PICVGLY 466  
DB 680 VSLILGRNANVNLSKSGTLPLHLAAGEDRVNVAEVLVNOGAHVDAQTGMGTPLHVCH 739  
QY 467 AOWGSK--SPLLK-KLEDEMKTFPGQOTPEPLFOFSWL--YVFLTLLOGGGLVFAFP- 520  
DB 740 --YGNIKIVFLLQHSKAVNAKKNQY--TPLHQAQOCHTIIINVLLON-----ASPN 790  
QY 521 ---VDTNLAIAISLSFLAIYIFVIVYEGGRREGESNMAMALSTRLARHIGYLELLEK 577  
DB 791 ELTVNNTALGI-----ARRIGYISVDT 814  
QY 578 LMFV-----NPPE-----LPEQTTKALPVRFLETDYNNLSV--GGET 613

```

Db      815 LKIVETTTTTEKHNMPETMNEVLMDSDDEVKANAPEMLSGEYISDVEGED 874
Qy      614 SIA-----EMIALSDACEREFGLATRL-----FRYFREESGKKKKKNOCL 658
Db      875 ANTGODKGLGPODLKELCDDSLPAEGTNGFSLGKRSLSFSDSRYTTLN--RSSAR 932
Qy      655 PEFVLEFLVGCITGILLALFVRDP- KHLT-----VNALISIASVGLAFVINC 709
Db      933 DSMMEELLVPSKEOHLFTFRREFDSDSLRHYSMADTLDNVNLVSSPIHSGFLVSFWDA 992
Qy      710 PFMVQVLDLSNOKRLHSAASKLHKLSEGMVVL-----KCEVELMARAKITDSTQ 765
Db      993 RG-----GSMGSRHHC-----MKIIPKRCIAP-----TR 1019
Qy      766 NOTRLVITIDGADCEODKVLQMLDTVRVLFSGKGFIAFASDPHIIKAINOMLSYL 825
Db      1020 ITCRLV-----KRKRLAN-----PPH-----GERRGISRLV 1047
Qy      826 DSNINGHYMRN- YHLEFVFLNSKGLSNARKFLVTSATNGDITGSDTGTGEDYDRYSQ 884
Db      1048 EMGPAGAOPLPVIVIEIPHFGSMRG--KERELIYLRSNGE-----TWREHOPD--SK 1096
Qy      885 NSLGEMTKL--GSKTALNRDYYRRQMRITITQMSFDLTKLVTEDEMEDISPOIMR 942
Db      1097 NE--DUTELNGMDELDSPELGKKRICKRITITDF-----POYFAVS--RIKQ 1142
Qy      943 LNIIVSYTGRLLRAN-----QITFMDRLASMINLTEQWPYRTSMLIYLETEGLPDM 997
Db      1143 ESNOIGPBGILSTYVPLVQASPEGALTKRINVLGQ-----APVDEI 1188
Qy      998 TLKTM-----YERISKIP-----TKKVEPLEI 1022
Db      1189 VKITIGKATSPYIVYVPRRRKFKPITMTIPIVPPSGEGSVNGYKDTPNRLCS 1248
Qy      1023 DG-----DIRNEFVLSRTFVLVARDVKFLPCGVNLDPLRELIANDVRAARQIN 1074
Db      1249 TGGSPAQWEDI-----TGTTLTFLTKDCVSF--TTNW--SARWILADCHOVLETVG 1296
Qy      1075 IGGIAYPPLPLHEGPPRPPSGYSOPASVCSASFNGPPGVSPOPHSSYYSGLSGPOH 1134
Db      1297 LATOLYREL-----ICV-----PYMAKFV-----VEAKMNDP-- 1323
Qy      1135 PFYRAAVPATGSSILSSMTVDVYVCEKRIQIEGLDNNMPOYTTIKKANINGRVLSQC 1194
Db      1324 -----VSSLRCPCMTDDVDVTLDEQE----- 1346
Qy      1195 NIDELKKEMANFGDMLFRSMVLEKRSVQVVPEDPRFLN-ENSSAPVPHGE-----S 1248
Db      1347 NFEV-----ARSKDIEVLEKGPYIVDCGNLAPLTKGQOLVEN 1386
Qy      1249 ARBSHTELP-----LTLESSQTPYTLNFSEBELNTGLDCAPRHNSLWSQOTRTPSL 1304
Db      1387 FYSFKENRLPFISIKIRIOTSEPCGRISFLKERKTKTGLPOTAVCNLNTITLPAHKETES- 1445
Qy      1305 SLSNSQDSSTEIKLTQKVOAEYRDAYRE--YIAQMSOLEGGTSSITIGRSSPHSTY 1361
Db      1446 -----DODDEIF--KTRROSFPASLAKRYSTLEGTMTERTGAT--RSLP--TTY- 1492
Qy      1362 IGOSSSGSIHSTLEOERGEGLKODGKRSFLMKRGDIVDYSSSGSVSTENASPLDPT 1421
Db      1493 -----SYKPFESTRPYQSWTAPITVGPAS-----GPTSLSSSSNTPSPASPLKSTW 1541
Qy      1422 EEDKSNQSGSKLLPGKSSERPSLYOTDLKLGGLIROKLPDSDEDESGTGAVQITPHC 1481
Db      1542 -----SVTSPSPKSTL-----GASTSPSVKSI SDV 1567
Qy      1482 SKMIRTKRLAKORECASPEHSAEPIKRFIKAKY-----LSDALLDK-----KXS 1528
Db      1568 ASPIRISLRT-----MSPIKTYVSSPINIOVSSGTLRAAPAVEATPLKGI 1614
Qy      1529 SDGCVRSNESSPNHSLHNEAADDSOLEKANLIELEDEHSGKRGMPHS----- 1576

```

```

Db      1615 ASNTESSRTSP-----VTTAGSLERSITMTPPASPKSNINMTSSSLPFKSIITSA 1668
Qy      1577 -----LSGIQDPIAMISICSEDKSPSECSLIASSPESWPA----- 1614
Db      1669 PLISSPILKSVSPKSVHVDIVISSAKITMAS--LSSPVQMGHAEVALVNGSISPLKY 1725
Qy      1615 COKAYNINRTPSTYTLNIN--TAPUNRANQ-----NPEIEIGIRETSQVI-----LRPSPN 1665
Db      1726 ASSSTLNGKATATLOEKISSATNSVSSVSAATQVEKVFSTTAMPFPLRSTVSA 1785
Qy      1666 PTAOVENLKSMAHKSQSSYTRLSKDSSELHAASSEST 1705
Db      1786 PSAFOSLTPSPAS-----ALYTSL--GSSISATTSVY 1816

RESULT 5
YB23_HUMAN
ID YB23_HUMAN STANDARD; PRT; 768 AA.
AC O9JL17;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA1223 (fragment).
GN KIAA1223.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
CC - SIMILARITY: CONTAINS AT LEAST 14 ANK REPEATS.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC CC
DR EMBL: AB033049; BAA86537.1; -.
DR HSSP: P42711; IDC2.
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ank. 14.
DR SMART: SM00248; ANK. 13.
DR PROSITE: PS50088; ANK. 13.
DR PROSITE: PS50297; ANK_REPEAT: 13.
KM Hypothetical protein; Repeat; ANK repeat.
FT NON_TER 1 1
FT REPEAT 1 11 ANK 1.
FT REPEAT 15 44 ANK 2.
FT REPEAT 48 82 ANK 3.
FT REPEAT 86 115 ANK 4.
FT REPEAT 119 148 ANK 5.
FT REPEAT 152 181 ANK 6.
FT REPEAT 185 214 ANK 7.
FT REPEAT 218 247 ANK 8.
FT REPEAT 251 280 ANK 9.
FT REPEAT 284 313 ANK 10.
FT REPEAT 317 346 ANK 11.
FT REPEAT 350 379 ANK 12.
FT REPEAT 383 412 ANK 13.
FT REPEAT 416 446 ANK 14.
FT SEQUENCE 768 AA; 82819 MW; 2913B69BE2DFE06D CMC64;

```





Db 490 DSKCLEYLRNDANPGI---RDQGYNAVHSAAYGHRICQLDIA-----SEPLDYLM 540  
 OY 441 YDLSSALADILSEPMOPPIGVGLAOWSGSKFLKLLED-EKKTAGQOTEPLOPS 499  
 Db 541 ETSIGDMSDSDRATIS-PLHLAAVHGHALELVYOSLDDVNRSSGRTPLDLAFAK 599  
 OY 500 WLIVPLTLILCGGLGV 516  
 Db 600 GHVECVDLINOGASIL 616

RESULT 7  
 DAPK\_HUMAN STANDARD; PRT; 1431 AA.  
 ID DAPK\_HUMAN  
 AC P53355;  
 DT 01-OCT-1996 (rel. 34, Created)  
 DT 01-NOV-1997 (rel. 35, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Death-associated protein kinase 1 (EC 2.7.1.-) (DAP kinase 1).  
 GN DAPK1 OR DAPK.  
 OS Homo sapiens (Human).  
 NC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBL\_TaxID-9606;  
 RN NCBI  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-95129831; PubMed-7828849;  
 RA Deiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;  
 RT Identification of a novel serine/threonine kinase and a novel 15-kD  
 RT protein as potential mediators of the gamma interferon-induced cell  
 RT death.  
 RL genes Dev. 9:15-30(1995).  
 RN [2]  
 RN REVISIONS TO 164-171.  
 RA Feinstein E.;  
 RL Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.  
 CC -1 FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL  
 CC DEATH.  
 CC -1 PFM: AUTOPHOSPHORYLATED.  
 CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -1 SIMILARITY: CONTAINS 10 ANK REPEATS.  
 CC -1 SIMILARITY: CONTAINS 1 DEATH DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 CC EMBL; X76104; CA53712.1; -  
 CC HSSP; Q63450; IAO6.  
 CC Genew; HGNC:2674; DAPK1.  
 DR DR  
 DR MIM: 600831; -  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000488; Death.  
 DR InterPro: IPR00719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam: PF00023; ank; 8.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00531; death; 1.  
 DR ProDom: PRO000001; Euk\_pkinase; 1.  
 DR SMART; SM00248; ANK; 7.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 6.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;

KW Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.  
 FT DOMAIN 13 266  
 FT DEATH 13 266  
 FT REPEAT 267 334  
 FT REPEAT 378 407  
 FT REPEAT 411 440  
 FT REPEAT 444 473  
 FT REPEAT 478 507  
 FT REPEAT 511 540  
 FT REPEAT 544 573  
 FT REPEAT 577 606  
 FT REPEAT 610 639  
 FT REPEAT 876 905  
 FT REPEAT 1163 1197  
 FT DOMAIN 1313 1397  
 FT NP\_BIND 19 27  
 FT BINDING 19 27  
 FT ACT\_SITE 139 139  
 FT MUTAGEN 42 42  
 SO SEQUENCE 1431 AA; 160017 MW; 9EB84811004A1558 CRC64;

Query Match 4.78; Score 414.5; DB 1; Length 1431;  
 Best Local Similarity 21.68; Pred. No. 5.5e-15;  
 Matches 221; Conservative 156; Mismatches 335; Indels 309; Gaps 36;

OY 5 ISOSVINYVEEENIPALKALECK--DVERNECGOPPLMAAEGQVVEIKLNGA 62  
 Db 344 YMAIITHAINDVNPGLIHLGLSLNDVQPKHGTPLPLAAGCGNLIQLIKRS 403  
 OY 63 NCNLEDNNTALISAKESGHIHVEELKSGASLEHRDGMGTALMAKCYGRDYEL 122  
 Db 404 RIDVDKGSNAYVMAARHGVDTLKFLSENKCPDLVKKSGEMALHVARGHADVAY 463  
 OY 123 LLSHGAMPSTGYLYSTYPTITMAAGRHADIVILLQNGAKVNSDKYGTPLVMAARG 182  
 Db 464 TCASAPIPISRTKEETPLHCAAMHYISVAKALCEACCNVINKREGETPLTASAG 523  
 OY 183 HLECVKHLAMGADVQEGANSMTALIVAKGYTSYVEILKRNPNVM/TPKDGNTALM 242  
 Db 524 YHDIVECLAHGAD-----LNAOCKDGHIALH 550  
 OY 243 IASKEGHEIVOLLDAAGTYVNPDSGDTLVIGARGHVEIVRALLQKADIDIRGD 302  
 Db 551 LAVRCOMETIKTLISQCGFVDYDRHGNTPLHVAACKDNMPYVALCE----- 599  
 OY 303 NKTALYAVEKGNATWVRDILQCPDTEICTKDGEPFLKATKRMRIEVELLDKAVY 362  
 Db 600 -----ANCNLD--ISKYGRPLHLAANGILDVYRLCLMGASY 637  
 OY 363 SAVDKGDTPLHVAIRGRSRRLAELLRNPKDGRLLYRPNKAGETPPYNIDCSHOKSILQ 422  
 Db 638 EALTDTGKTAEDLARSEQHENVAGLLARLKD-----THRGLEIQQ 678  
 OY 423 IFGARHLSPTENDGMLGTDIVSSALADILSEPTMQPPICVGLYAWMGSGKFLKLLED 482  
 Db 679 -----LRPQ-----NLQPRIKLFLHSGSGKTTLVESLK- 709  
 OY 483 EMKTFAGQOTEPLOFSWLIVETLLICGGLV-----AFVDYTNLA 526  
 Db 710 -----CGLLASFRRRRRSLSTNSSRFPSPPLASKPT 742  
 OY 527 IAIISLFLALYIFPVIYIFGRRBGSNNMAMALSTRLARHIGYIELFLKIMVNPPEL 586  
 Db 743 VSVSINN-----YPCENSVYRSRSMPEPGLTK--GMLE-----VEVAPTHH 784  
 OY 587 P-----EORTKAL-----PYRPLFTDYNKLSSVSGFTSLAEMLA 620  
 Db 785 PHCSADQSTKADIONATLVNVDGFSWERSGNPVYFCYD--FANDPSTSHVVF 841  
 OY 621 TISDACEREG--FLATRLFRVFRTES--QGRKKWKTCCLPSPYIFLYVGCIIAG 674  
 Db 842 SLEEPYEQLNVPVIFWLSFKSLVVEEPLIAGGALK-----NPLQYVL----- 885  
 KW 675 ITLLAIFRVDPKHLVYNNALISIASVGLAFVLANCRITWQVYLDLSINSQRR--LHSA 730



Query Match 4.5%; Score 395.5; DB 1; Length 1327;  
 Best Local Similarity 23.3%; Pred. No. 5.5e-14;  
 Matches 167; Conservative 80; Mismatches 211; Indels 259; Gaps 17;

14 EENIPALKALEKCKDVERNEGOTPLMLAEOGNEYVELLKNKANCLLEDLDMWT 73  
 348 EKKIMALLPLVNVNCHASGRK---STPLHLAAGNRRIRAYOLLQHGADVNAKKGGLV 404  
 74 ALISAKBGHIIIVEELLKSGASLEHRDNGMTALMAYCGRDVEVLLSHGANPSV- 132  
 405 PLHNCYGHYEVEVELLKHGACVNAWMDMOTPLHEAASKRVEVCSILLSHGADPTLV 464  
 133 -----T 133  
 465 NCHGSAVDMAFPELREKRYTEFKGSHLLQARADLAKYKKTALLETINFKOPQSHET 524  
 134 GLQYSV-----YPLIMAGRGHADIHVLHLONGAKV 164  
 525 ALHCAVASLHPKRRKQVTELLRKGANVKNKDFEPLHVAERAHNDVMEVLHKGAKM 584  
 165 NCSDKYTPPLVMARKHLECVKLLMAGADVDOEGANSMTA----- 207  
 585 NALDTLGGTALRALALHLCFCRLLSYSDPSIISLQGTAAOMGNEAVOOLISESTP 644  
 208 -----LIVAVAGYQSYKEILKRNPNVNLTKDG--NTALIMASKEGHETIYODL 256  
 645 IRTSDVRYLEASAGLETKOCS--SONVNCDELEGRSTPLHFAAGNRRSVYEVYL 703  
 257 LDAGTYVNPBRSGTGLVAVRGHVEIVRALLQKVDIDIRGDNKTALYMAVEKG-- 314  
 704 LHHGADVNAKDKGLVPLHNCYGHYEVEVELLVHGAASVNAVDLMKFTPLHEAANKGY 763  
 315 -----NATVNDILO-----C 325  
 764 EICKLLKHGADPTKKNNDGNTPLDYKEGDTIDDLKGAALLDAKKCCLARVOKLC 823  
 326 NEDTICTKD--GETPLIKATKRNIEVELLDKGAASVANDKGDTPPLH----- 374  
 824 TPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAOKGGLIPLHNAASYGHV 883  
 375 -----VAIRGRSRRLALLNRKDGGLLYRPNKAGETP 408  
 884 DIALLLKYVNCVNAQTMATPPLHEAOKGRT-OLCLALLHAGADPMT--KNQECOTP 939  
 409 YN-----IDCSHOKSILT-----OIFGARHLSPTETDGMGLYDLSALADI 451  
 940 LDLATADDIRALLIDAMPPEALPTCFKQATVVASLSISPASTPS-----CLSAASSTDN 994  
 452 LSEPTMOPPIVGYLAOMGSG-----KSFLLKLEDEMKT 486  
 995 LTPPLAE--LAVGGASNAGDAAGTERKEGEVAGLDMNISQFLSLGLEHLRDIFFET 1049

RESULT 9  
 ASB2\_HUMAN STANDARD; PRT; 587 AA.  
 ID ASB2\_HUMAN STANDARD; PRT; 587 AA.  
 AC Q96Q27; Q9NSU5; Q9Y567;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ankyrin repeat and SOCS box containing protein 2 (ASB-2).  
 GN ASB2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND INDUCTION.  
 RX MEDLINE=21450655; PubMed=1156180;  
 RA Kohroki J., Fujita S., Itoh N., Yamada Y., Imai H., Yumoto N.,  
 RA Nakanishi T., Tanaka K.;  
 RA "ATRA-regulated Asb-2 gene induced in differentiation of HL-60

RT leukemia cells.";  
 RL FEBS Lett. 505:223-228(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND INDUCTION.  
 RX MEDLINE=21634875; PubMed=11682484;  
 RA Gubal F.C., Moog-Lutz C., Smolewski P., Di Giola Y.,  
 RA Dazynkiewicz Z., Lutz P.G., Cayre Y.E.;  
 RT "ASB-2 inhibits growth and promotes commitment in myeloid leukemia  
 cells.";  
 RN J. Biol. Chem. 277:218-224(2002).  
 RL [3]  
 RP SEQUENCE OF 31-587 FROM N.A.  
 RC Tissue-uterus;  
 RA Oltensaeider B., Obermaier B., Mewes H.-W., Gassenhuber J.,  
 RA Wiemann S.;  
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 42-587 FROM N.A.  
 RX MEDLINE=20564172; PubMed=11111040;  
 RA Klie B.T., Viney E.M., Willison T.A., Broditch T.C., Cancilla M.R.,  
 RA Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,  
 RA Alexander W.S.;  
 RT "Cloning and characterization of the genes encoding the ankyrin repeat  
 and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";  
 RL Gene 258:31-41(2000).  
 CC -1- FUNCTION: Not known. Its expression induces growth inhibition and  
 chromatin condensation.  
 CC -1- INDUCTION: By all-trans retinoic acid (ATRA).  
 CC -1- SIMILARITY: CONTAINS 12 ANK REPEATS.  
 CC -1- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB056723; BAB64532.1; -  
 EMBL; AJ251238; CAC17765.1; -  
 EMBL; AL137735; CAB70899.1; -  
 EMBL; AF159164; AAD45345.1; -  
 Genew; HGNC:16012; ASB2.  
 MIM; 605759; -  
 DR Interpro: IPR002110; ANK.  
 DR Interpro: IPR001496; SOCS.  
 DR Pfam: PF00023; ank; 11.  
 DR PROSITE; PS50088; ANK\_REPEAT; 9.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50225; SOCS; 1.  
 KW ANK repeat; Repeat.  
 FT REPEAT 56 85  
 FT REPEAT 89 119 ANK 1.  
 FT REPEAT 123 152 ANK 2.  
 FT REPEAT 156 185 ANK 3.  
 FT REPEAT 189 218 ANK 4.  
 FT REPEAT 222 251 ANK 5.  
 FT REPEAT 255 284 ANK 6.  
 FT REPEAT 288 317 ANK 7.  
 FT REPEAT 320 349 ANK 8.  
 FT REPEAT 362 391 ANK 9.  
 FT REPEAT 392 421 ANK 10.  
 FT REPEAT 428 456 ANK 11.  
 FT REPEAT 453 587 ANK 12.  
 FT DOMAIN 533 587 SOCS BOX.  
 FT CONFLICT 224 224 MISSING (IN REF. 4).  
 FT CONFLICT 460 460 MISSING (IN REF. 4).  
 SQ SEQUENCE 587 AA; 65083 MW; 866A2B6010C4C1F9 CRC64;

Query Match 4.3%; Score 384.5; DB 1; Length 587;  
 Best Local Similarity 30.4%; Pred. No. 6.2e-14;  
 Matches 113; Conservative 71; Mismatches 161; Indels 27; Gaps 6;

QY VINVEBENIPALKALEKCKDYDERNEGOTPLMLAEOGNEYIVEKLLK-NGANCNLE 67  
 Db 61 LIRAIRKGDDEALKTKMKREKNLAEPKKEMLPHEAAYGYOGCKLVORALPGITIDOR 120  
 QY 68 DLDMDALLISASKEGHIEIHELKSGASLEHRDGMGTALMMACYKGRDVELLSHG 127  
 Db 121 TLDEETAVLATCGHIDCLLSLQACAEBDISNKSRETPLYKACERKNAEVAIVIOHN 180  
 QY 128 A-----NPSVTGQGYVYPIYMAAGRGADYIHLLOKNAKNCSTYKTTPLVMAARK 181  
 Db 181 ADTNHRCNRGWTALHESV-----SRNDLEVMQILVSGAKESKNAKYITPLFYAAQS 233  
 QY 182 GHLECEVHLLAMGADVOEGANSMTALIVAKGYTOSVEILKRNPNVMLTRKDGNTAL 241  
 Db 234 GQLEALFLAKYGDITFOASDNASALYEACKNHEVEFLLSQGADAKTKNDGLPL 293  
 QY 242 MIASKEGHEIIVODLDAGTYVNIIPDRSGDTVLIGAVRGHVEITRALLQKVDID- 297  
 Db 294 HIAKSKGNRYIVQMLLPVTSRRI-RRSGVSPHLIAERNHDEVLALISARFVNTPLA 352  
 QY 298 -----INGDNKTKALYNAVEKGNATWRIQCPRDEICTKDEPFLIKATKRNIEVY 352  
 Db 353 PERALYEDRRSSALYFVANNVYATELLQHGADP---NRDVISPLVAIRHGCLRTM 409  
 QY 353 ELLDKGAKVSA 364  
 Db 410 QLLDHGANIDA 421

RESULT 10  
 TNR2\_HUMAN STANDARD: PRT; 1166 AA.  
 AC 09H2K2: 09HAS4: 09H8F2:  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TRF1-  
 DE interacting ankyrin-related ADP-ribose polymerase 2) (Tankyrase-1like  
 DE protein) (Tankyrase-related protein).  
 OS TANK2 OR TANKL OR TANK2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP TISSUE: Fetal brain;  
 RC MEDLINE=21072350; PubMed=11205898;  
 RA Monz D., Munnia A., Comtesse N., Fischer U., Steudel W.-I., Feiden W.,  
 RA Glaes B., Meese E.U.;  
 RT "Novel tankyrase-related gene detected with meningioma-specific  
 RT sera.";  
 RL Clin. Cancer Res. 7:113-119(2001).  
 RN [2]  
 RP TISSUE: Fetal brain;  
 RC TISSUE: Breast carcinoma;  
 RA MEDLINE=21190090; PubMed=11294570;  
 RA Kurosh A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K.,  
 RA Scallan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;  
 RT "Cloning and characterization of TANKL, a member of tankyrase gene  
 RT family.";  
 RL Genes Immun. 2:52-55(2001).  
 RN [3]  
 RP TISSUE: Liver;  
 RC MEDLINE=21264473; PubMed=11278563;  
 RA Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M.,  
 RA Elyre H.J., Sutherland G.R., Daly R.J.;  
 RT "Identification of a novel human tankyrase through its interaction  
 RT with the adaptor protein Gbl4.";  
 RL J. Biol. Chem. 276:17172-17180(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.

RC TISSUE=Placenta;  
 RX MEDLINE=21443728; PubMed=11454873;  
 RA Kaminker P.G., Kim S.-H., Taylor R.D., Zebrafidian Y., Funk W.D.,  
 RA Morin G.B., Yasven P., Campisi J.;  
 RT "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes  
 RT rapid induction of cell death upon overexpression.";  
 RL J. Biol. Chem. 276:35891-35899(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND  
 RP LNEP/OTASE.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=21661461; PubMed=11802774;  
 RA Shodio J.I., Lodish H.F., Chi N.-W.;  
 RT "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1  
 RT (telomere-repeat-binding factor 1) and IRAP (insulin-responsive  
 RT aminopeptidase).";  
 RL Biochem. J. 361:451-459(2002).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Yin Y., Gelmann E.P.;  
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Tracey A.;  
 RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE OF 838-1151 FROM N.A.  
 RC TISSUE=Placenta;  
 RA Tsogal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fuji A., Hara H.,  
 RA Tanase T., Nomura Y., Toshiya S., Komai F., Hara R., Takeuchi K.,  
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
 RA Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Oshima A.;  
 RT "NPO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP FUNCTION, AND ADP-RIBOSYLATION.  
 RX MEDLINE=21602874; PubMed=11739745;  
 RA Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;  
 RT "Role for the related poly(ADP-ribose) polymerases tankyrase 1 and 2  
 RT at human telomeres.";  
 RL Mol. Cell. Biol. 22:332-342(2002).  
 CC -1- FUNCTION: May regulate vesicle trafficking and modulate the  
 CC subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP  
 CC activity and can modify TRF1, and thereby contribute to the  
 CC regulation of telomere length.  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-riboseyl)(N)-acceptor =  
 CC nicotinamide + (ADP-D-riboseyl)(N+1)-acceptor.  
 CC -1- SUBUNIT: Oligomerizes and associates with TANKS. Interacts with the  
 CC cytoplasmic domain of LNEP/OTase in SLC2A4/GLUT4-vesicles. Binds  
 CC to the N-terminus of Gbl4 and TRF1 with its ankyrin repeat  
 CC region.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and  
 CC with juxtanuclear SLC2A4/GLUT4-vesicles. Also found around the  
 CC pericentriolar matrix of mitotic centrioles. During interphase, a  
 CC small fraction of TANK2 is found in the nucleus, associated with  
 CC TRF1.  
 CC -1- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,  
 CC liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral  
 CC blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,  
 CC ovary, small intestine, colon, mammary gland, breast and breast  
 CC carcinoma, and in common-type meningioma. Highly expressed in  
 CC fetal liver, heart and brain.  
 CC -1- PTM: ADP-ribosylated (-auto).  
 CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -1- SIMILARITY: CONTAINS 15 ANK REPEATS.  
 CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial



FT REPEAT 229 258 ANK 7.  
 FT REPEAT 262 291 ANK 8.  
 SQ SEQUENCE 323 AA; 33453 MW; 3976D3A49BD07DF CRC64;

Query Match 4.38; Score 378.5; DB 1; Length 323;  
 Best Local Similarity 36.1%; Pred. No. 5.2e-14;  
 Matches 97; Conservative 43; Mismatches 128; Indels 1; Gaps 1;

QY 24 LLEKCDVDERNCGOTPLMLAEGQNVETIKELKNCNEDDNTALISAKESH 83  
 DB 51 LLEQGAHVDPNPGFKTALFAVEGDDLETVALLSRGDVNAKRYAGCTPLFEAECH 110  
 QY 84 IHIVEELKSGASILEHRDMGWTALMACYKGRDVEELLISGANPSYGLQSYPTII 143  
 DB 111 IGISALLERGAHVHRTSGMDALIASRHGTTDMVEQLLEFGADPKAADREGRT-ALM 169  
 QY 144 MAARGHADIVHLLONGAKVNCSDYGTTPYMAARKHGLECVKHLAMGADVDOEGAN 203  
 DB 170 QAAKSGETGYPLLLIGSGADLEARDQGAATALLIADQAGAVETIAGIGADLDVADL 229  
 QY 204 SMALIVAVKGYTQSVKEIKRNPVNLTKDQNTALIMASKEGHEIYODLDAGTYV 263  
 DB 230 GSTALLIAGHGVAVVERLLANGADPNRODRKGTALMEAVTDHAEIDRLIAGART 289  
 QY 264 NIDRSGDVTLLIGAVGSHVEIYRALLQK 292  
 DB 290 DLKDDADRTADIAVQUGHQAIARLAPK 318

RESULT 12  
 ANR3\_HUMAN STANDARD; PRT; 832 AA.

AC P57078: Q96KH0;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2002 (Rel. 41, Last annotation update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat domain protein 3) (PKC-delta-interacting protein kinase).  
 GN ANKRD3 OR DIK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Fetal kidney, and Fetal lung;  
 RA Shimizu N., Kudoh J., Shibuya K.;  
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; Pubmed=10830953;  
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Saitani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordstieck G., Hornischer K., Brandt P., Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Blocker H., Ramser J., Beck A., Klages S., Hennig S., Risselmann L., Dagnold E., Wehrach H., Reinhardt R., Yaspo M.-L.;  
 RA "The DNA sequence of human chromosome 21.";  
 RT Nature 405:311-319(2000).

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -1- SIMILARITY: CONTAINS 10 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcements/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AB047783; BAB56136.1;  
 CC EMBL: AP001743; BAA95526.1;  
 CC HSSP: P25963; 1IKN.  
 CC Genew: HGNC:496; ANKRD3.  
 DR MIM: 605705; ANK.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR004040; STY\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR01245; Tyr\_kinase.  
 DR Pfam: PF00023; ank; 10.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_kinase; 1.  
 DR SMART: SM00248; ANK; 10.  
 DR SMART: SM00221; STYK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE: PS50088; ANK\_REPEAT; 9.  
 KW Transferase: Serine/threonine-protein kinase; ATP-binding; Repeat; ANK repeat; Alternative splicing.

FT DOMAIN 22 286  
 FT REPEAT 485 514 ANK 1.  
 FT REPEAT 518 547 ANK 2.  
 FT REPEAT 551 580 ANK 3.  
 FT REPEAT 584 613 ANK 4.  
 FT REPEAT 617 647 ANK 5.  
 FT REPEAT 651 680 ANK 6.  
 FT REPEAT 684 713 ANK 7.  
 FT REPEAT 717 746 ANK 8.  
 FT REPEAT 750 780 ANK 9.  
 FT REPEAT 782 811 ANK 10.  
 FT NP\_BIND 28 36 ATP (BY SIMILARITY).  
 FT BINDING 51 51 ATP (BY SIMILARITY).  
 FT ACT\_SITE 143 143 BY SIMILARITY.  
 FT VARSPIC 278 325 MISSING (IN ISOFORM 2).  
 FT CONFLICT 714 714 M -> V (IN REF. 1).  
 SQ SEQUENCE 832 AA; 91610 MW; 5D8FFD5F047ECB CRC64;

Query Match 4.1%; Score 360; DB 1; Length 832;  
 Best Local Similarity 27.5%; Pred. No. 2.5e-12;  
 Matches 109; Conservative 60; Mismatches 126; Indels 102; Gaps 6;

QY 29 KVDENRNECGOTPLMLAEGQNVETIKELKNCNEDDNTALISAKESH 66  
 DB 477 QVVDALDSGASLHLVADAGEECACKMLNNANPNLSNRGSPPLMAVERRYGYVE 536  
 QY 67 -----BDLNTWALISAKESHIVEELKSGASILEHRDMGWTALMACYK 115  
 DB 537 LILAKRISVANADEQWTALHFAQNGESSRLLEKNASVNEDEFGRRPMHVAQCHG 596  
 QY 116 RTDVEYELLISGANPSYGLQSYPTIYMAAGRGADIVHLL-ONGAKVNCSDYGTTP 174  
 DB 597 QENIVRILLIRRGVDSLOG-KDAMVLP LHHAAQGHLPYKLLAKPGSVNAQTLDGTRP 655  
 QY 175 LMAARKHGLECVKHLAMGADVDOEGANSMALIVAVKGYTQSVKEIKRNPVNLTD 234  
 DB 656 LHLAAGRGHVRARILLIDCSVNCSLAQPLHVAEFTGHTSARILLIRGAGKEAMT 715  
 QY 235 KDGNTALIMASKEGHEIYQDLDAGTYVINIPDRSGDVTLLIGAVGSHVEIYRALLQK 294  
 DB 716 SDGYTALHLAA-----RNGHLATYKLLVEEKA 742  
 QY 295 DIDIDGQDKTALIVAVKGNATWTRDILQCNPTIECTKQCEPLIKATKRNIEVEL 354  
 DB 743 DVLARGLPMTALHLAAAGHSEVEIEV-----SADVIDL 778



OY	355 LIDGAKVSAVDKKGDTPLHVAIRGSRRIAELLRN 391   :   :   :   :   :   :   :
Dd	779 F-----DEQGSALHLAAQGKAQIVETLLRH 805

AC	DT	ID	RA	STANDARD	PRT	741 AA.
AC	005823;	AC	RNSA_HUMAN			
DT	01-NOV-1995 (Rel. 32, Created)					
DT	01-FEB-1996 (Rel. 33, Last sequence update)					
DT	15-JUN-2002 (Rel. 41, Last annotation update)					
DE	2.5A-dependent ribonuclease (EC 3.1.26.-) (2'-5A-dependent RNase)					
DE	(Ribonuclease L) (RNase L) (Ribonuclease 4).					
GN	RNASEL OR RNSA.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OC	NCBI_TaxID=9606;					
RA	[1]					
RA	SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-240 AND LYS-274.					
RA	TISSUE=Kidney.					
RA	MEDLINE=93201598; PubMed=7680985;					
RT	Zhou A., Hassel B.A., Silverman R.H.;					
RT	"Expression cloning of 2'-5A-dependent RNase: a uniquely regulated					
RT	mediator of interferon action.";					
RT	Cell 72:753-765(1993).					
RT	[2]					
RT	SEQUENCE FROM N.A.					
RT	TISSUE=Placenta.					
RT	MEDLINE=20515260; PubMed=11063255;					
RT	Zhou A., Nie H., Silverman R.H.;					
RT	"Analysis and origins of the human and mouse RNase L genes: mediators					
RT	of interferon action.";					
RT	Mamm. Genome 11:989-992(2000).					
RT	[3]					
RT	CHARACTERIZATION OF RNASEL ACTIVITY.					
RT	MEDLINE=94245737; PubMed=7514601.					
RT	Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,					
RT	Silverman R.H.;					
RT	"Intrinsic molecular activities of the interferon-induced 2'-5A-					
RT	dependent RNase.";					
RT	J. Biol. Chem. 269:14153-14158(1994).					
RT	[4]					
RT	FUNCTION, AND SUBCELLULAR LOCATION.					
RT	PubMed=11585831;					
RT	Le Roy F., Bisbal C., Silhol M., Martinand C., Lebou B.,					
RT	Salenazda T.;					
RT	"The 2'-5A/RNase L/RNase L inhibitor (RNI) pathway regulates					
RT	mitochondrial mRNAs stability in interferon alpha-treated H9 cells.";					
RT	J. Biol. Chem. 276:48473-48482(2001).					
RT	[5]					
RT	ERRATUM.					
RT	Le Roy F., Bisbal C., Silhol M., Martinand C., Lebou B.,					
RT	Salenazda T.;					
RT	J. Biol. Chem. 277:13354-13354(2002).					
RT	[6]					
RT	REVIEW.					
RT	PubMed=9856285;					
RT	Castelli J., Wood K.A., Youle R.J.;					
RT	"The 2'-5A system in viral infection and apoptosis.";					
RT	Biomed. Pharmacother. 52:386-390(1998).					
RT	[7]					
RT	MUTAGENESIS OF LYS-392.					
RT	PubMed=9862963;					
RT	Dong B., Silverman R.H.;					
RT	"Alternative function of a protein kinase homology domain in 2',					
RT	5'-oligoadenylate dependent RNase L.";					
RT	Nucleic Acids Res. 27:439-445(1999).					
RT	[8]					
RT	MUTAGENESIS OF HIS-583; PRO-584; TRP-632; ASP-661; ARG-667 AND					
RT	HIS-672.					

```

FT DOMAIN 229 242 2-5A BINDING (P-LOOP) 1.
FT DOMAIN 253 275 2-5A BINDING (P-LOOP) 2.
FT ZN_FING 395 444 C6-TYPE (ATYPICAL).
FT VARIANT 59 59 G-> S.
FT VARIANT 406 406 /FID-VAR_013509.
FT VARIANT 406 406 S-> F.
FT VARIANT 462 462 /FID-VAR_013510.
FT VARIANT 462 462 R-> O (RISK FACTOR FOR HPC1; IN
FT DBSNP:486907).
FT VARIANT 541 541 /FID-VAR_012056.
FT VARIANT 541 541 D-> E (IN DBSNP:627928).
FT MUTAGEN 240 240 K->N: REDUCED 2-5A BINDING ACTIVITY;
FT MUTAGEN 240 240 ALMOST COMPLETE LOSS OF 2-5A BINDING
FT MUTAGEN 274 274 ACTIVITY WHEN ASSOCIATED WITH N-274.
FT MUTAGEN 274 274 K->N: REDUCED 2-5A BINDING ACTIVITY;
FT MUTAGEN 392 392 ALMOST COMPLETE LOSS OF 2-5A BINDING
FT MUTAGEN 392 392 ACTIVITY WHEN ASSOCIATED WITH N-240.
FT MUTAGEN 392 392 K->R: COMPLETE LOSS OF ENZYMACTIC ACTIVITY
AND ENZYME DIMERIZATION. NO CHANGE IN
FT MUTAGEN 583 583 BINDING TO 2-5A AND RNA.
FT MUTAGEN 583 583 H->A: NO CHANGE IN ENZYMACTIC ACTIVITY.
FT MUTAGEN 584 584 P->A: NO CHANGE IN ENZYMACTIC ACTIVITY.
FT MUTAGEN 632 632 W->A: NO CHANGE IN ENZYMACTIC ACTIVITY.
FT MUTAGEN 661 661 D->A: COMPLETE LOSS OF ENZYMACTIC
ACTIVITY.
FT MUTAGEN 667 667 R->A: COMPLETE LOSS OF ENZYMACTIC
ACTIVITY. NO CHANGE IN 2-5A BINDING AND
FT MUTAGEN 672 672 ENZYME DIMERIZATION.
FT MUTAGEN 672 672 H->A: COMPLETE LOSS OF ENZYMACTIC
ACTIVITY. NO CHANGE IN 2-5A BINDING
FT SEQUENCE 741 AA; 83532 MW; 91365EA307E3CELD CRC64;
ACTIVITY AND ENZYME DIMERIZATION.
Query Match 4.08; Score 357.5; DB 1; Length 741;
Best Local Similarity 27.48; Pred. No. 2.8e-12;
Matches 121; Conservative 64; Mismatches 128; Indels 129; Gaps 13;

```

```

ID RN5A_MOUSE STANDARD: PRT; 735 AA.
AC 005921; QPERUT;
AC 01-NOV-1995 (rel. 32, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE 2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN RNASEL OR RN5A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C3H/An; TISSUE=adipose tissue;
RX MEDLINE=20515260; PUBMED=11063255;
RA Zhou A., Nie H., Silverman R.H.;
RT "Analysis and origins of the human and mouse RNase L genes: mediators
of interferon action."
RL Mamm. Genome 11:989-992(2000).
RN [2]
RN SEQUENCE OF 1-679 FROM N.A.
RX MEDLINE=93201598; PUBMED=7680958;
RA Zhou A., Hassel B.A., Silverman R.H.;
RT "Expression cloning of 2-5A-dependent RNase: a uniquely regulated
mediator of interferon action."
RL Cell 72:753-765(1993).
RN [3]
RN FUNCTION, AND SUBCELLULAR LOCATION.
RP PubMed=11585831;
RX Le Roy F., Bisbal C., Silhol M., Martinand C., Lebou B.,
RA Salehzada T.;
RL "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
mitochondrial mRNA stability in interferon alpha-treated H9 cells."
RT J. Biol. Chem. 276:48473-48482(2001).
RN [4]
RN ERRATUM.
RA Le Roy F., Bisbal C., Silhol M., Martinand C., Lebou B.,
RA Salehzada T.;
RL J. Biol. Chem. 277:13354-13354(2002).
CC -1- FUNCTION: Endoribonuclease, mediator of interferon action, which
play a role in mediating resistance to virus infection and
apoptosis. Might play a central role in the regulation of mRNA
turnover.
CC -1- CATALYTIC ACTIVITY: Cleaves 3' of UPNP dimers, with preference for
UU and UA sequences, to sets of discrete products ranging from
between 4 and 22 nucleotides in length.
CC -1- COFACTOR: Optimal RNA cleavage rates requires the presence of
either manganese or magnesium and ATP.
CC -1- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
linked oligoadenylates) the homodimerization and subsequent
activation occurs. Inhibited by RNase L inhibitor.
CC -1- SUBUNIT: Monomer (inactive form) or homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC -1- TISSUE SPECIFICITY: Expressed in spleen, thymus, lung, testis,
kidney, liver and heart.
CC -1- INDUCTION: By interferons.
CC -1- DOMAIN: A single active nuclease domain in a dimer is
sufficient for ribonuclease activity.
CC -1- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
constitute the 2-5A binding domain.
CC -1- SIMILARITY: CONTAINS 9 ANK REPEATS.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; AF281045; MAG33708.1; -
DR

```

DR EMBL: L10382; AAA37117.1; -  
 DR HSSP: P42773; 11H8.  
 DR MGD; MG1:1098272; Rnasel.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00023; ank; 8.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 8.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE: PS50088; ANK\_REPEAT; 7.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc-finger;  
 KW Repeat; ANK repeat.  
 FT REPEAT 24 53 ANK 1.  
 FT REPEAT 58 87 ANK 2.  
 FT REPEAT 91 120 ANK 3.  
 FT REPEAT 124 153 ANK 4.  
 FT REPEAT 167 197 ANK 5.  
 FT REPEAT 201 234 ANK 6.  
 FT REPEAT 238 268 ANK 7.  
 FT REPEAT 272 301 ANK 8.  
 FT REPEAT 303 328 ANK 9.  
 FT DOMAIN 364 384 PROTEIN KINASE LIKE.  
 FT DOMAIN 585 735 RIBONUCLEASE.  
 FT DOMAIN 229 242 2-5A BINDING (P-LOOP) 1.  
 FT DOMAIN 253 275 2-5A BINDING (P-LOOP) 2.  
 FT ZN\_FING 401 436 C6-type (POTENTIAL).  
 FT SEQUENCE 735 AA; 83274 MW; B6632F4A5B50F711 CnC64;

Query Match 4.0%; Score 356; DB 1; Length 735;  
 Best Local Similarity 32.2%; Pred. No. 3.4e-12;  
 Matches 110; Conservative 57; Mismatches 111; Indels 64; Gaps 11;

QY 73 TALISASKEGHIHYEELKSGALEH-RDMGSGTALMAGCYKRTYVELLSHGANS 131  
 DB 27 SSLTKAOKGVVAVVQDLEKGAJANACEDTWTPLHNAVQAGRVIVLLSHGADP 85  
 QY 132 VTGLQSYVPIYMAAGRHADIVHLLONGAKVNCSDKYGTPLVMAARKGHLECYVHL 191  
 DB 86 -----HRRKKNGA-----TPITAGIGDVKLELIL 112  
 QY 192 AMGADVDOEGANSMTALIVAVKGGYTSYKEILKRNPNVL---TDKD-----GNTAL 241  
 DB 113 SCGADVNECDENGFTAFMEAAERGNABALRFLPAKGANVNLRRQTTDKRRLKOGGATL 172  
 QY 242 MIAKEGHIIEVOLL-DAGTYVNIIPRSGDYVILGAVRGHV---EIVRALLQYADI 296  
 DB 173 MSAAEKGLHEVLRLLNDKRAEVDARNMGRNALIRTLWDCENVEITSILOHGDV 232  
 QY 297 DIRGQDKTALYVAVKGNATVVDIQQ---CNPDTICTKDGSEPLIKATKRNIEVE 353  
 DB 233 NVREGRKTPLIANAVERKHGIVOMLSREGINIDAR--DNEGKTALLIADVOKLEIYO 290  
 QY 354 LLLDKGAKVSAVDKGDTPLEHVALIGRSRLAEILL---RNP 392  
 DB 291 LLEKGA-----DKCDLWIAIRRNHDYHLVKLLPYANP 326

## RESULT 15

ANR6\_HUMAN STANDARD; PRT; 692 AA.  
 AC Q9Y2G4; Q9Y2G4; Q9Y2G4; Q9Y2G4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ankyrin repeat domain protein 6.  
 GN ANKRD6 OR KIAA0957.  
 OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=99246063; PubMed=10231032;  
 RA Nagase T., Ishikawa K.-I., Suyama K., Kikuno R., Hirose S.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.  
 RT "Prediction of the coding sequences of unidentified human genes. XIII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 6:63-70(1999).  
 [2]  
 RN SEQUENCE OF 87-692 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Testis;  
 RA Pouster A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE OF 140-692 FROM N.A. (ISOFORM 1).  
 RP Tracey A.;  
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 [4]  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE  
 CC -1- PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 8 ANK REPEATS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AB023174; BA76801.1; -  
 DR EMBL: AL117504; CAB55968.1; -  
 DR EMBL: AL096678; CAB86658.1; -  
 DR HSSP: Q00420; 1AWC.  
 DR GeneW; HGNC:17280; ANKRD6.  
 DR InterPro: IPR002110; ANK.  
 DR Pfam: PF00023; ank; 8.  
 DR PRINTS: PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 6.  
 DR PROSITE: PS50088; ANK\_REPEAT; 6.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 KW Repeat; ANK repeat; Alternative splicing.  
 FT REPEAT 9 38 ANK 1.  
 FT REPEAT 41 70 ANK 2.  
 FT REPEAT 74 103 ANK 3.  
 FT REPEAT 107 136 ANK 4.  
 FT REPEAT 140 169 ANK 5.  
 FT REPEAT 173 202 ANK 6.  
 FT REPEAT 206 235 ANK 7.  
 FT REPEAT 239 268 ANK 8.  
 FT VARSPLIC 264 264 ANK 8.  
 FT Q -> OVLFSEKSLRKKRRLKEERRAQSVPDEVAQSK  
 FT (IN ISOFORM 2).  
 FT N -> I (IN REF. 2).  
 FT CONFLICT 87 87  
 FT SEQUENCE 692 AA; 75722 MW; D32115487A80839F CnC64;

Query Match 3.9%; Score 349.5; DB 1; Length 692;  
 Best Local Similarity 29.2%; Pred. No. 7e-12;  
 Matches 120; Conservative 67; Mismatches 177; Indels 47; Gaps 11;

QY 95 ASLEHRDMGWTALMAGCYKRTYVELLSHGANSVTGLQSYVPIYMAAGRHADIV 154  
 DB 8 AALSER-----LLVAYKQGTENVYQILNKGARVAVT--KHRTPLHLAANKGHLPV 58  
 QY 155 HLLONGAKVNCSDKYGTPLVMAARKGHLECYVHLLAMGADVDOEGANSMTALIVAVK 214  
 DB 59 QILAKAGCDLVYQDDGDTALIRATVVGNTETIALIHGECALDRQKDGDTALHESW 118  
 QY 215 GYTQSVKILKRNPNVNLTDGNTALMIAKEGHIIEVOLL-DAGTYVNIIPRSGTFL 274

```

Db      119 GFSQSAKLLVKAGANYLAKNKAGNTLHLACQNSHSQSTRVLLLAGSRADLKNNAGDTCL 178
QY      275 IGAVRGHEIIVRALLQKYADIDIRGQDNKTALYWAWEKGNATWVRDIIQCNPDTETICTK 334
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      179 HVAARYNHLSIIRLLLTAFCSVHEKQAGDTALHVAALNHHKKVAKILLEAGADTTIVNN 238
QY      335 DGETPLIKATKMNIEVVELLDKGAKYSVNDKGDTP-LHVAIRGRSRRLAELLRNPK 393
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      239 AGQTPLETARYHNNPEVA-LILTKAPQGSV--SAGDTPSSQAVARKEAREEFLSASPE 295
QY      394 ----DGRLLYRPNKAG---ETPYNIDCSHOKSILTOIFGARHLSPTETD----- 435
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      296 PRAKDDRRRRSRKRYSAFSDPTPPADQDPGHQKNL-----HANNHPKKRNHRHRCSSPP 349
QY      436 -GDMLGDIYSSALADILSEPTWQPPICVGLYAQWGSKSFLLKKLEDEMK 485
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      350 PHEFRAYQOLYT--LYRGKDKGWQAPIN-----GCRCPEPLINKLENOLE 391

```

Search completed: July 1, 2003, 14:35:02  
 Job time : 33 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:05 ; Search time 43 Seconds  
(without alignments)  
3834.198 Million cell updates/sec

Title: US-10-021-571-2  
Perfect score: 8853  
Sequence: 1 MSVLISQSVINVEENIPAL.....ELHAASSTGTGGERESIL 1715

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5073.5	57.3	1031	2 T43458	hypothetical prote
2	1848.5	20.9	1398	2 T21884	hypothetical prote
3	1024	11.6	246	2 T46392	hypothetical prote
4	555	6.3	1881	1 S4HUK	ankyrin 1, erythro
5	553	6.3	426	2 AE2149	hypothetical prote
6	553	6.2	1856	2 B35049	ankyrin 1, erythro
7	553	6.2	1880	2 A35049	ankyrin 1, erythro
8	552	6.2	1961	2 T42716	ankyrin 3, splice
9	551.5	6.2	1940	2 T42715	ankyrin 3, splice
10	551	6.2	1943	2 T42713	ankyrin 3, splice
11	550	6.2	3924	2 S37431	ankyrin 2, neurona
12	527.5	6.0	1862	2 I49502	ankyrin - mouse
13	526	5.9	4377	2 A55575	ankyrin 3, long sp
14	516.5	5.8	1848	2 S37771	ankyrin, erythrocy
15	510.5	5.8	1765	2 T42714	ankyrin 3, splice
16	479	5.4	2039	2 T15347	ankyrin-related un
17	464.5	5.2	1549	2 T13940	ankyrin - fruit fl
18	429	4.8	2584	2 T24158	hypothetical prote
19	429	4.8	2606	2 T24157	hypothetical prote
20	414.5	4.7	1423	1 T37275	death-associated p
21	409.5	4.6	791	2 T42691	hypothetical prote
22	397	4.5	1062	2 T30255	inversin - mouse
23	384.5	4.3	557	2 T46507	hypothetical prote
24	382	4.3	1062	2 T14151	inv protein - mous
25	378.5	4.3	323	2 B47169	ankyrin-like repea
26	361	4.1	397	2 T46445	hypothetical prote
27	356	4.0	679	2 B45771	2-5a-dependent RNA
28	347.5	3.9	741	2 A45771	2-5a-dependent RNA
29	338.5	3.8	991	2 T25412	hypothetical prote

30	335	3.8	1401	2 S11527	alpha-latrotoxin p
31	325	3.7	1188	2 T19552	hypothetical prote
32	323.5	3.6	441	2 S74626	erythroid ankyrin
33	321.5	3.6	934	1 H1274	probable ankyrin -
34	302	3.4	1058	2 D82654	ankyrin-like prote
35	301	3.4	2004	2 D88948	protein ZK1005.1 f
36	300.5	3.4	1411	2 S30355	alpha-latrotoxin
37	296.5	3.3	627	2 E86190	hypothetical prote
38	294	3.3	1184	2 T00253	gene Ankhzn protei
39	292	3.3	606	2 AC2508	hypothetical prote
40	288	3.3	1083	2 S48460	probable membrane
41	287	3.2	517	2 T48283	ankyrin-like prote
42	286.5	3.2	662	2 E84725	ankyrin-like prote
43	285	3.2	1001	2 S30385	G9a protein - huma
44	279.5	3.2	368	2 T18184	ankyrin repeat pro
45	279	3.2	543	2 C86212	hypothetical prote

## ALIGNMENTS

RESULT 1  
T43458  
hypothetical protein DKFZP434F0621.1 - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000  
C:Accession: T43458  
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
Submitted to the Protein Sequence Database, December 1999  
A:Reference number: 222517  
A:Accession: T43458  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1031 <AAA>  
A:Cross-references: EMBL:AL133620  
A:Experimental source: adult testis; clone DKFZP434F0621  
A:Genetics:  
A>Note: DKFZP434F0621.1

Query Match	Similarity	Score	DB 2	Length
Best Local Similarity	95.2%	Pred. No. 4.8e-293		
Matches	974	Conservative	13	Indels 13; Gaps 1;
QY	1	MSVLISQSVINVEENIPALKLEKCKVDENEGGPTPLMAAGNVEIKELKN	60	
DB	1	MSVLISQSVINVEENIPALKLEKCKVDENEGGPTPLMAAGNVEIKELKN	60	
QY	61	GANCLEEDLDNMTALISAKKEGHIHYEELKSGASLEHRDMGWTALMAACYKGRDYY	120	
DB	61	GANCLEEDLDNMTALISAKKEGHIHYEELKSGASLEHRDMGWTALMAACYKGRDYY	120	
QY	121	ELLISHGANPSVTGLQTSYVPIIYAAGRGADYIHLLONGAKYCSDKTTPPLVVAAR	180	
DB	121	ELLISHGANPSVTGLQTSYVPIIYAAGRGADYIHLLONGAKYCSDKTTPPLVVAAR	180	
QY	181	KGHEECVYKHLAMGADVDOEGANSMTALIYAVKGYTOSYEILKRPNVNLDPKDGNTA	240	
DB	181	KGHEECVYKHLAMGADVDOEGANSMTALIYAVKGYTOSYEILKRPNVNLDPKDGNTA	240	
QY	241	LMISAKKEGHIYVDLLDAGTYVNIIPRSGDPTVLIGAVRGHVEIVALLQKYADIDIRG	300	
DB	241	LMISAKKEGHIYVDLLDAGTYVNIIPRSGDPTVLIGAVRGHVEIVALLQKYADIDIRG	300	
QY	301	ODNKTALYMAVEKGNATVVRRIIDCNPTETCTDGETPLIKAKMNIIEVELLDKGA	360	
DB	301	ODNKTALYMAVEKGNATVVRRIIDCNPTETCTDGETPLIKAKMNIIEVELLDKGA	360	
QY	361	KVSADVKKGDTPLVAIIRGSRRLAELLRNPKGRILYPRNKAGEPPYNDISHOKSIL	420	
DB	361	KVSADVKKGDTPLVAIIRGSRRLAELLRNPKGRILYPRNKAGEPPYNDISHOKSIL	420	
QY	421	TOIFGARHLSPETEDGMDLGYDLYSSALADILSEPTMQPICVGLYVQMGSSGSFLLKLL	480	
DB	421	TOIFGARHLSPETEDGMDLGYDLYSSALADILSEPTMQPICVGLYVQMGSSGSFLLKLL	480	









	F:601-633/Domain:	ankyrin repeat	homology <AN18>	
	F:634-666/Domain:	ankyrin repeat	homology <AN19>	
	F:667-699/Domain:	ankyrin repeat	homology <AN20>	
	F:700-732/Domain:	ankyrin repeat	homology <AN21>	
	F:733-765/Domain:	ankyrin repeat	homology <AN22>	
	F:766-798/Domain:	ankyrin repeat	homology <AN23>	
Query Match	6.2%; Score 553;	DB 2;	Length 1856;	
Best Local Similarity	20.5%; Pred. No.	5.3e-24;		
Matches 388;	Conservative 287;	Mismatches 648;	Indels 574;	Gaps 73;
QY	21 LKALLECKKDDEENEGGPTPLMAAEAGNVEIKELKNGCANONLEDJOWTALLISASK	80		
Dd	94 VRELVTGANVNAASQKGFTPLMAADENHLEVKFLLENGANOVATEDGFPLAVALQ	153		
OY	81 EGH-----	83		
Dd	154 QGHENVVAHLINYGTKGVRLPALHIAARNDPRTAAVLLONDNPDLSTGTGTPPLHIA	213		
QY	84 -----IHIVELLKSGASLEHRDGGWTALMACYKGRFDVVELLSHGANSPTYGLQYS	138		
Dd	214 AHENLVANQLLLNGSSVNFTPPONGIITPLHIASRGNVIWRLIDRGAOIE-TKTKE	272		
QY	139 VYPIITMAAGRHADIVILLONGAKVNCSDKYTGTFPLVMARKHGLECYKKLLMGADVD	198		
Dd	273 LTPLHCARNGHVRISELTLDHGAPTOAKTRNGESPIHMAAQGHLDLCVRLILDYDAEID	332		
QY	199 -----ORGA-----NSMTALVAVVGGYTSVKELK	225		
Dd	333 DITTDHTLPPLVAACHGHRAVAKYLIDGGAAPSBNALNGFPLHIACKKNHYRVRELLK	392		

```

Db 926 MGRSRHN-----GLRVVLP-----PRTCAAPTRITCRV----- 954
QY 783 DKVQLMDLVVRLVFSKGPRIALFASDPRIILAIKAIKONLSVLRDSNINSHDMRN-IYHL 841
Db 955 --KPOLSTPPPL-----AEEGGLASRIALGPACAFSLSVIYEI 993
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITCSDTGTGODDRRSNLSIGEMTIGSKTALN 900
Db 994 PPHASHGRG---DRELVLVRSNGSV-----WKEHRSRGESYIDQIILN-GMDELIG 1041
QY 901 RRDYTRRQMOQTTRQKSFDTKLIVT---EDWESDISPQTRRLINIVSYGLLRAN 957
Db 1042 SLEELKEKRVCKRIIT--TDFPLYFVIMSLCOD-YDILCEG-----GSLKSLVPLV 1091
QY 958 QITFPMMDRLASMINLTEQWPRYSMLILYLETBGLPQOMTLK-----TWER 1005
Db 1092 QATFPEPNVTKKVKALQ-----AQVPPDELVTKLIGNATSPVITVEPR 1137
QY 1006 ISK-----NIPTKDVEBLEIDGDIRNEFV-----LSSRTPLVA 1042
Db 1138 RKKFHRPIGLRIPLPSPMTDNPRDGEEDTSLRLCSVIGGTDAQMEDITGTKLVA 1197
QY 1043 ROKVTFPLCTVLDKRLRITADVAAARQINIGLAPPLPLHGGPPPSGCSQASV 1102
Db 1198 NECANF---TTNV--SARFWLSDCPRTAEAVNFATLTKEL-----TAVPYMAKF 1242
QY 1103 CSSASFNGPFCGVVSPQPHSSYSGLSQPHFYNRAVPATSGSLSSMTVYVCEK 1162
Db 1243 VIFAKMNDPREG-----RLCYCMTDKVUKT 1269
QY 1163 LKQIGLQONMPPQYCTTK---ANINGRVLSCNIDELKEMANFGDWHLERS--M 1216
Db 1270 LEQHNFEVARSRIEVLGMSLFAELSG-----NIVPYKAAQGRSPHQSPRENRL 1323
QY 1217 VLEMSVSEQVVP-----EDPRFLNENSAVP-----HEBSARRSHHTLP 1258
Db 1324 AMPVAVRDSRPPGSGLSFLKAMKYEDTQHLCLNTMPCACGSAEDRRRTPTPLA 1383
QY 1259 L--TELSQTPYTLNFSFE-ELNTLGLDEGAPRHSNLSWOSQTRTP--SLSLNSQDSSI 1314
Db 1384 LKYSILSESTPGSLGTDQAEKMAVISE---HLGSMALARELQPSVDIN----- 1433
QY 1315 ELSKLDKVAEYRRAVREYIAQMSQLEGGSSTISGRSSPHSTYIT--GSSSGGSIH 1372
Db 1434 -----RIRVENPMSLLEQSVALLNTL-----WVIRCONANMENIX 1468
QY 1373 STLEO-ERKGEKELQEDGRKSFILMK---RGVDYIDSSGYSTNEASPLDITTEDE--- 1425
Db 1469 TALQSIDRGEIYVNMLEGGSGROSRNLKPRRHTDRYSLSPSOMNGYSSL-----QDELLS 1523
QY 1426 -----KSDQSGSKL-----LPKKSSEPSLFTQDLKLGGLRYOKLPSPDED 1468
Db 1524 PASLICALSPPLRAQYVNEVAAILDAIPLAETHMTLEMSDMQVWSAGLT--PSLVAED 1582
QY 1469 ESGTRVQITPHCSKMITRKLAKQRCASPOEHSAPITPTIKAKYILSDALLDKDS 1528
Db 1583 SS-----LECSK---AEDSDATGHEMKLEGALSEPRPELGSLTEVDDVDSAT 1631
QY 1529 S-----DSGVRNESPNSHLSHNEAADDSOLEKANLILEDEGHSGKGMPSHLSGL 1580
Db 1632 NGLIDLQEGEGOREEKLPGSKRQDDATGAGQ--DSENEVSLVSGHQGRQARITHS--- 1686
QY 1581 ODPITARMSICSEDK-----KSPSECSLIASSPEBSW 1612
Db 1687 --PTVSQVTERSODRLQDMDADGSIYSYLQDAAGSW 1721

```

## RESULT 7

A35049  
 ankryn 1, erythrocyte splice form 2 - human  
 N:Alternative names: ankryn 2.1, erythrocyte; ankryn-R  
 C:Species: Homo sapiens (man)

```

C:Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C:Accession: A35049
R:Lambert, S.; Yu, H.; Prechal, J.T.; Lawler, J.; Ruff, P.; Spelcher, D.; Cheung, M.C.
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankryn.
A:Reference number: A35049; MUID:90175370; PMID:1669849
A:Accession: A35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1880 <LAMB>
A:Cross-references: GB:M28880
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: Bp11.2-bp11.2
C:Superfamily: ankryn; ankryn repeat homology
C:Keywords: alternative splicing; cytoskeleton
F:2-1880/Product: ankryn 1, erythrocyte form 2 #status predicted <MA>
F:2-1513,1676-1880/Product: ankryn 2.2, erythrocyte #status predicted <MA2>
F:44-76/Domain: ankryn repeat homology <AN01>
F:77-109/Domain: ankryn repeat homology <AN02>
F:110-142/Domain: ankryn repeat homology <AN03>
F:143-171/Domain: ankryn repeat homology <AN04>
F:172-204/Domain: ankryn repeat homology <AN05>
F:205-237/Domain: ankryn repeat homology <AN06>
F:238-270/Domain: ankryn repeat homology <AN07>
F:271-303/Domain: ankryn repeat homology <AN08>
F:304-336/Domain: ankryn repeat homology <AN09>
F:337-369/Domain: ankryn repeat homology <AN10>
F:370-402/Domain: ankryn repeat homology <AN11>
F:403-435/Domain: ankryn repeat homology <AN12>
F:436-468/Domain: ankryn repeat homology <AN13>
F:469-501/Domain: ankryn repeat homology <AN14>
F:502-534/Domain: ankryn repeat homology <AN15>
F:535-567/Domain: ankryn repeat homology <AN16>
F:568-600/Domain: ankryn repeat homology <AN17>
F:601-633/Domain: ankryn repeat homology <AN18>
F:634-666/Domain: ankryn repeat homology <AN19>
F:667-699/Domain: ankryn repeat homology <AN20>
F:700-732/Domain: ankryn repeat homology <AN21>
F:733-765/Domain: ankryn repeat homology <AN22>
F:766-798/Domain: ankryn repeat homology <AN23>

```

Query Match 6.2%; Score 553; DB 2; Length 1880;  
 Best Local Similarity 20.5%; Pred. No. 5.4e-24;  
 Matches 388; Conservative 287; Mismatches 648; Indels 574; Gaps 73;

```

QY 21 LKALECKYDERNECGTPTMLAABQGNVEYVELKLNKANGCNLEDDNNTALISAK 80
Db 94 VRELVTNGAVNVAOSQGFPLVMAOENHLEVKFLENGANQNVATEDGFTPLAVALQ 153
QY 81 EGH----- 83
Db 154 QGHENNVYAHILNTGTGKVRPLALHIAARDRTAAVLQNDPNVDYLSKGTPTPLHIA 213
QY 84 -----IHYVELLSGASLEHRDMGWTALMAVCYGRDYVELLSHGANSPTGLQYS 138
Db 214 AHYENLVNVAQILNRRSSVFTPGNGITPLHIASRGNVYIWRLLDRQAQLE--TKTDE 272
QY 139 VYPTIYAAGRGHADIYHLLQNGAKVNCSDKYTTPPLVAAKRGHIECYKHLIAGADV 198
Db 273 LTPHCAARNGHVRISIELLDHGAPIQARTKGLSPIHMAAOGDHLDVRYLLQYDAEID 332
QY 199 -----QEGA-----NSMTALIVAKGYTQSVKEILK 225
Db 333 DITLDHLPLHVAHGHGHRVAKVLLDKAKRPSRLNFTPLHIAKKNHVRVWELLIK 392
QY 226 RNPVNLTJDKDGTALMALASKGHEIYVDLDACTYVNI PDRSGDYVLIGAVRGHVEI 285
Db 393 TGA SIDAVESGILPLHVASFMGHLPYVKNLLQRGASPVVSNVKTETPLHMAARAGHTEV 452
QY 286 VRALLQRYADIDIRQDNKTALYWAVERKNATVWRDILQCPDTEICTKDGFTPLIKATK 345

```

```

Db 453 AKYLLONKAVAKAKDDQPLHCAARIGHTNMVKLENNANPNLATTAGHTPLHIAAR 512
QY 346 MKNIEVELLDKGAQVAVDKKGTPLHVAIRGSRRLAELLNPNPDGLLYPNKAG 405
Db 513 EGHVEVIALLEKEASQACMTKKGTPLHVAIRKRVAVELLLE--RDA--HPNAG 566
QY 406 E--TPYNIDCSHO--KSL--TOFGAR-- 427
Db 567 KNGPLPLHVAHHNNLDIYKLLPRGSGSPHANNNGYTPPLHIAAKONOEVARSLLOXG 626
QY 428 -----HSPETDGMGLGYDLYSSALADISEPTMOP-----PIC-- 462
Db 627 SANAESVGVTPPLHIAAGHAEMVALLSLKQANGKSGLTPLHVAAGHGVADV 686
QY 463 -----VGLYAOMSGK--SFLKLEED--EMKTFAGQGTPLPOFSL 501
Db 687 LKHGVMDATRTMGVTPPLHVAHYGNIKLVKFLLOHQADVNAKTRLG--SPLHQAAQ 744
QY 502 --IVFTLLCGGLVFAFPVD-----TNLAISLSFLALITYFFVIYFGGRGE 553
Db 745 GHRTDYTLKKG-----ASPNEVSDGTPPLAIAKRLGYISYTVLKVY----- 789
QY 554 SWMAWALSTRLAHIGYLLFKLMFVNPPELPGQTKALPVRELTIDYRLS--SVGE 612
Db 790 TDETSFVLSDKH--MSPEYDEIDVSEDEGEELIS-----FKAERDSRDVDE 840
QY 613 TSLAEMIAFLSDACEHEFGFLATRLR-----VFTEF--SQCKKKMKTKCCLPSY 662
Db 841 KELLDFVPKIDOVE-----SPALPRICAMPEYVIRSEOEQAKREYDSDSLPS-- 892
QY 663 IFLEIVGCIAGITLAIFRVDPKHLTVNAIISIVAGLAFLNCRFMQVLDLSLNS 722
Db 893 -----SPATISDNI--SPVAPVHTGFLVSF-----MVDARGS 925
QY 723 QRRRLHSAASKLHLKESGFMKYLCEVELAMAMAKTIDSFQONQTRLVIIIDGLDACEQ 782
Db 926 MRGSRHN-----GLRVYIP-----PRCAAPRITCRVY----- 954
QY 783 DKYLOMLDVRVLFSGPFLAIPASDPHIIKAINONLSVLRDSINHDYMRN--IYHL 841
Db 955 --KPKQLSTPPPL-----AEEGLASRIITALGPTGAQFLSPYIVEI 993
QY 842 PVFL--NSRGLSNARKFLVTSATNGDITCSDTTGTOTDRVRSQNSIGEMTKLSGTALN 900
Db 994 PHRASGRG--DRELVLVLRSEGSV-----KHEHSRGESLTDQILN--GMDEELG 1041
QY 901 RRDYRRRQMRQRTTQMOSFDLTKLVT--EDWESDISPQMRRLINVTYTGRLIRAN 957
Db 1042 STEELKKKRCRIT--TDFPLFYVIMSRLCD--YDIIGPEG-----GSLKSKLPLV 1091
QY 958 QITFNMRLASWINLEOMYRTSWILVLENEGLPDQMTLK-----TMYER 1005
Db 1092 QATFPENAVYKRYKLLQ-----AQPVPDELVTYKLLCONATFSEPTIVEBP 1137
QY 1006 ISK-----NIPTKDVEPILLEIDGIRNFEV-----LSSRPVLVA 1042
Db 1138 RRFHHRIGIRILPSPWMDNPRDSGDTSLRLCSVIGTDQAWEDITGTTLVYA 1197
QY 1043 RDVKTPLCTVNDPKIREIIADVRAAREOINTGATAPPLPLHCEGPPRPSPSY 1102
Db 1198 NECANF--TTNV--SARFMLSDCPTAEAVNFATLLYKEL-----TAVPYMAKF 1242
QY 1103 CSSASFNGPFGVSPQPHSSVYSGLSGPHFYRAAVPATGSSILLSSMVDVACEK 1162
Db 1243 VIFAKMNDPREG-----RLRCYCMDDKVDYKT 1269
QY 1163 LROIIGLDQMMPOYCTTIK--ANINGRVLSCNIDELKEMANFEGDWHLFRS--M 1216
Db 1270 LEQHEHNEVEARSDIEVLEGMSLFAELSG-----NLVYVKAQAQORSFHFSPRENRL 1323
QY 1217 VLEKRSVESOVVP-----EDPRFLNENSAPVP-----HGEARSSSHTELP 1258
Db 1324 AMPYKVDSSRREGSLSFLRKAKMYEDTOHIIHLNITMPKCAKSGAGDRRTPTPLA 1383

```

```

QY 1259 L--TELSQPYTLNFSPE--ELNTLGLDEGAPRHSNLSMOSQTRRP--SLSSINSODSSI 1314
Db 1384 LRYLSLSESTIGSLSGEQAKMAVISE-----HLGSLMELARELOFVEDIN----- 1433
QY 1315 EISKLPKVOAEYDAYREYIAOMSOLEGGTSSFTIGRSSPHSTYI--GSSSGGSIH 1372
Db 1434 -----RIVENPNLSLEQSVALLN-----WVIREQONANMENLY 1468
QY 1373 STLEQ--ERKGEGLKOEDEGRKSFMLK--RGDVIDYSSGVSTNEASPLDPIREDE-- 1425
Db 1469 TALQSIDGELVNNLEGGSGRSNMLKPPDRRTDSDYLSPSQMGYSYL-----QDELS 1523
QY 1426 -----KSDQSGSLK-----LPGKSSERPISFTDLKLGKGLRYOKLPSEDE 1468
Db 1524 PASLQALSSPLRADQYNEVAIIDAIPLATIETHDTHLEMSQDMVMSAGLT--PSLYTAED 1582
QY 1469 ESGTGRVOITPHCSKMITRKLKAKKORCASPOEHSAPRTITIKAEYSDALLDKDS 1528
Db 1583 SS-----LECSK--AEDSDATGHEWKLEGALSEPRGPELSLEVEDDVTDSDAT 1631
QY 1529 S-----DSGVSNSSSPHNSLHNEAADSOLEKANLIELEDEBHGSKRMHSLSG 1580
Db 1632 NGLIDLBOEGGOREEKLPGSKRODDATGAGQ--DSENEVLSVGHORGARITHS---- 1686
QY 1581 QDPIARMSICSEDK-----KSPSECSLIASSPEESW 1612
Db 1687 --PTVSQYTERSQDRLQMDADGISTVYIQLQAAQSGW 1721

```

## RESULT 8

T42716  
 ankyrin 3, splice form 4 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
 C:Accession: T42716  
 R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo J.; Cell Biol. 130, 313-330, 1995  
 A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene repeat domain.  
 A:Reference number: 222237; MUID:95340633; PMID:7615634  
 A:Accession: T42716  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1961 <PEP>  
 A:Cross-references: EMBL:U40632; NID:6710548; PID:6710552; PIDN:AAB01607.1  
 A:Experimental source: strain C57BL/6J; kidney  
 A:Gene: Ank3  
 C:Genetics:  
 A:Map position: 10  
 C:Superfamily: ankyrin repeat homology  
 C:Keywords: alternative splicing

Query Match 6.2%; Score 552; DB 2; Length 1961;  
 Best Local Similarity 20.2%; Pred. No. 6.7e-24;  
 Matches 427; Conservative 289; Mismatches 674; Indels 724; Gaps 85;

```

QY 14 EENPIPAKALLEKKDVERNEGOTPLMLAEOGNEIYKELKNGANCNLEDDIWMY 73
Db 66 KEGHEVVSSELLQREANDATAATKGNATLHLSAGQEVYKAVLYTNCAVNAASQNGFT 125
QY 74 ALISAKEGHIIHVELLKSGLSLEHRDMSGNTALMAACYGRNDVVELL----- 124
Db 126 PLYMAAQENHLEVNFLLDNGASQSLATBEDGTPPLAVALQOCHDOVLSLELNDTKGKVR 185
QY 125 -----SIGANPSYTG- 135
Db 186 LPALHIAKDDTKAAILLQNDTNADVESKGTPLHIAHYGININATLLNRAAAND 245
QY 136 --OVSVPPIWAAGGHADIYHLLONGAKVNSDKYGTTPPLVWAARKG----- 182
Db 246 FTARNDITPLHVASKRGANMVKLLLDKGAIKADATRGITPLHGAASGHQVYEMLLD 305

```

QY	18	-----	-----	-----	HECVHLLAMGADVDOEGANSFTALIVANKGYTOS	213
Db	306	RSAPILSKKNGLSPLMAHTQGDHLCNVOLLIQHNVPDYDTNDYLTALHVAHCGHYKV	-----	-----	-----	365
QY	220	VKEILKRNPNVNLTDKGNLTALMASKE	-----	-----	-----	247
Db	366	ACVLLDKKSPAKNALNGFTPLHACKNRIRMYELLKKGASIGAVTSSGLPRHYAAF	-----	-----	-----	425
QY	248	-GHIEIYODLLAGFVYVNIIDRSQDVLIGAVRGHVEIVARLLQKYADIDRGDNKTA	306	-----	-----	-----
Db	426	MCHVATVSOILMHGASPTTNRGFEALTMAARSGOAEVVRKYLVODGAEVAKKADDDYR	485	-----	-----	-----
QY	307	IYWAVEKGNATVVRDILQCNPDTEICTKDEPFLIKATMRNIEVVELLDGAKASYAD	366	-----	-----	-----
Db	486	LHISARLKGADIVOOLOOGASPMNATTSGYRPLHAAEGHEVAAFLDHBGASLSTT	545	-----	-----	-----
QY	367	KKGDPRLHVAHNGRSRRLLAEILLR	-----	-----	NPKGRLL	398
Db	546	KKGFTPLHAAVYKGLIEVASILLQKSASPDAGKSGLPRHYAAHYDNOKVALLLLDQGA	605	-----	-----	-----
QY	399	--YRPNKAGETPYNIDC--SHOKSILTOI--FGAR-----	-----	-----	-----	-----
Db	606	SPHAAKNQYRPLHIAAKKNQMDIATSLLEYADANNVYRQSIASVHLAAQGHYDMSL	665	-----	-----	-----
QY	442	DLXSALADILSEPTMQP-----	-----	-----	-----	-----
Db	666	LISRANANVLSKSGELTPLHAAQDVRVNAEVLVNOGAHYDAQTGMGYPLHVGCH--Y	723	-----	-----	-----
QY	470	SGSK--SFLK---KLEDEMKTFAGQTEPLEFOFSMLVYFLV-----	-----	-----	LLGGLGLV	516
Db	724	GNIKIVNFLLOHSAVNAKTKMGYATALQAQOGHTHIIINVLLQNNASPNELTVNG	779	-----	-----	-----
QY	517	FAFPDVLTAIATISFLAIYIFPIV-	-----	-----	-----	544
Db	780	-----NTALAIARRGYISVVDVLEIWTTEITTEKHKNMNPETMNEVLDSDDEV	834	-----	-----	-----
QY	545	-----YFGGRREGESNMAMALSTRLARHIGYLELLKL	-----	-----	-----	582
Db	835	RKASPERKLSDEYISDDEGDKCWFKIPKYQ-----	-----	-----	EVLKSEDAITGDTDKLG	886
QY	563	PPELPEQTTKALPVRFLEFTDNRKSSVGGETSLAMIAITLSDACEREGFLATRLFRYR	642	-----	-----	-----
Db	887	PODLKELDDSDILPAE-----GYVG--FSLGARSASLR-----	-----	-----	SESSDRSYTLNR	928
QY	643	TEESQKKKKKKKTCOLPSFVIFLFIYVGLINGITLALFRPDR-KHL-----	-----	-----	-----	693
Db	929	SSYAR-----DSMMTEELVPSKEOHLFTTFREDSDSLHXYMAADDTLDNMVIL	977	-----	-----	-----
QY	694	LISIASVYGLAFVLCRTMQVLDLSLSQ-----	-----	-----	RLHSAASKLHKL	737
Db	978	SSPVHSGFLVSPMDARG-----GSMRGRHGHKIIIPRKCTAPRTYITGLYKRNHLA	1032	-----	-----	-----
QY	738	-----KSEG-----	-----	-----	-----	766
Db	1033	NPPEVBEGLASRLVEMGPAGAOFLGPIVEIPIHFGSMRKERELILRSENGETWEKH	1092	-----	-----	-----
QY	767	Q-----TRLVYIIGDLADCEQDKVLOMDYTRV--LFSKGFLIAFASDPHIIIIKAIKON	819	-----	-----	-----
Db	1093	QFDSKNEDLALINGMD--EELDSPELGTRKRICHIITK-DFOYFA-----	-----	-----	-----	1144
QY	820	LNSVLRDSNNGHDHMYRNIVHLPVFLNSRGLSNARKFLV--TSATNGDITCSDDTGTQ--	875	-----	-----	-----
Db	1145	SNOIGPEBGI-----	-----	-----	-----	-----
QY	876	--EDPDRVSONSLGEMTKGSKTALNRDYY--RRQMQTTIRQMSFDLTKL--	-----	-----	-----	928
Db	1185	VPEETVKKI-----	-----	-----	-----	-----
QY	929	EDWFSIDSPQMRRLNLIVSYTG-----	-----	-----	-----	980
Db	1233	NGYKGDATPNL--RLT--CSITGTSAPQWEDICTGPTPLPIKNCVSPITTVNSARF	1284	-----	-----	-----
QY	981	SML-----ILVLETEBGLPDMQMLKTM-YEISKNIPPTKQDVEPBLEITGDGRNF	1029	-----	-----	-----







QY 1599 SECS--LIASPEEWACQAKAYNLNRPSTVTLNNTAPTRANONPEDEIG 1649  
 Db 1872 EFTSKYIEDAPKPCVPVGMKK-----TRTADGKARLNQDEEG 1912

## RESULT 10

T42713  
 ankyrin 3, splice form 1 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
 C:Accession: T42713  
 R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Vialamas, M.; Turtzo, J.  
 J. Cell Biol. 130, 313-330, 1995  
 A:Title: Ankyrin (epithelial ankyrin), a widely distributed new member of the ankyrin gene  
 the repeat domain.  
 A:Reference number: 222237; MUID:95340633; PMID:7615634  
 A:Accession: T42713  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1943 <PEP>  
 A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1  
 A:Experimental source: strain C57BL/6J; kidney  
 A:Genetics:  
 A:Gene: Ank3  
 A:Map position: 10  
 A:Introns: 855/1  
 C:Function:  
 A:Description: supposed to play an important role in the polarized distribution of many  
 C:Superfamily: ankyrin; ankyrin repeat homology  
 C:Keywords: alternative splicing

Query Match 6.2% Score 551; DB 2; Length 1943;  
 Best local similarity 20.2%; Pred. No. 7.5e-24;  
 Matches 426; Conservative 289; Mismatches 666; Indels 724; Gaps 85;

QY 14 EENETPALKALIECKDVERNEGQTPMLAEOGNEIYKELKNGANCLEDDIMWT 73  
 Db 66 KEGHEVVESELLQREANVDATKRGNTALHIASLAGAEVVKVLTNGANVNAOSQNET 125  
 QY 74 ALISAKBGHIIYELLKSGASLEHRMGCTALMAACYGRDVEYELL----- 124  
 Db 126 PLYMAOENHLEVEFLDNGASOSLATEDEFTPLAVALQGHQVYSLLENDTKGVR 185  
 QY 125 -----SHGANPSVTGL----- 135  
 Db 186 LPALHTAARKDDTKAAALLQNDINADVESKSGFTPLHIAHYNINATILLRAAAYD 245  
 QY 136 ---QTSVPTIYAGRGHADIYHLLONGAKVNCSDKYGTPPLVWAARKG----- 182  
 Db 246 FTARNDITPLHVASKRGANVVKLLDRGAKIDAKTRDGLPLHCGARSGHEQVEMLLD 305  
 QY 183 -----HLECYKHLLANGADVDOGSASMTALLIYAVNGGYTOS 219  
 Db 306 RSAPILSKTKNGSLPLHATQGDHNCVQLLDHNNVPDVTNDYLTALHVAACHGVK 365  
 QY 220 VKEILKRNPNVNLTDKDGNTALMTASKE----- 247  
 Db 366 AKVLLDKKASPNAKALNGFTPLHACKKNRIYWEILLKHGASTOAVTESGLTPIHYAF 425  
 QY 248 -GHEIYVDDLDAGTYNIPRSGDYVLIGAVRGHEIYRALLQKYADIDIRGDNKTA 306  
 Db 426 MGHVNIYSQLMHNHGAASPTWTNVRGETALHMAARSGOAEVYRIVQDAQVAKAKDQTP 485  
 QY 307 LYMAVEGNATVBDIIQCNPDTEICRKDGTEPLIKATKMNIEYVELLDKGAAYAVD 366  
 Db 486 LHIAIRLCKADIVQQLQOGASPNAAATSGTPLHIAARSGHEDVAAFTLLDHGASISITT 545  
 QY 367 KKGDTPHVAIRGRSRRLAELLR-----NPKDGRLL----- 398  
 Db 546 KKGTPLHVAAYKGLKLEVASLLQKASPAAGSGGLTPLHVAHAHYNOVALLILLDOGA 605

QY 399 ---YRPNKAGETPNVINDC-SHOKSILNOI---FGAR-----HLSPETDGMGLY 441  
 Db 606 SPHAAKNAGTTPLEIAAKKNOMDIATSLLEYGADANAVTROGLASVHLAAQEGHVMVSL 665  
 QY 442 DLYSSALADILSEPTMOP-----PICVGLYAQW 469  
 Db 666 LLSRNANVNLSNKGSLPLHIAQEDRVNAEVLVNGAHVDAQTKMGYPLHVGCH--Y 723  
 QY 470 GSGK--SFLLK---KLEDEMKTRAGQOETPLFQFSWLYVFTL-----LLCGGLGLV 516  
 Db 724 GNIRIVFLLQHSKAVNAKTRNGYTALHQAQOGHTIIVLLQNNASPNELVNG----- 779  
 QY 517 FAFVVDNMLIAISLALITYFEIYI----- 544  
 Db 780 ---NIALAARLGLIISYVDILKVTYEELIMTTTTEKKNVPEYMEVLDMSDEX 834  
 QY 545 -----YFGRGEGESNMWMAALSTRILHIGYELLKIMFVNPELPEQOT 591  
 Db 835 RKASAPKLSIDGEIYSDGEED-----AITGDTDKYLQ-----PQDLKELGD 877  
 QY 592 KALPVRFLFDYNRSLSSVGETSLAEIATLSDACEREFGLATRLFVRFVEESGKKK 651  
 Db 878 DSLPAE-----GYVG--FSIGARSASLR-----SFSSDRSYTLNRSYSAR----- 915  
 QY 652 WKTCCLPSFVIFLEIYGCIIAGITLLAIFRVPD-KHLT-----VNAIISIASVYG 702  
 Db 916 ---DSMIEELVPSKEQHILTFREDSDSLRHSVADTLNVLVSSPVASGL 968  
 QY 703 LAFVLCRTWMOYVLSILNSQ-----RK-----RLHSAASKLHLK-----KSE 740  
 Db 969 VSEFVVDARG-----GSMRGRHHGMRIIPRCKTAPRTLCYLRHKLANPVPVVEGE 1023  
 QY 741 G-----FMKYLKVELAMAMATIDISFTONO-----TFL 770  
 Db 1024 GLASRLVEKMPAGAQLFQPIYVEIPHFGSRGRERELIYLRSENGETWKEHODSKNEDL 1083  
 QY 771 VVIIDGADCEQDKVLOMLDTPVR--LFSKGPPIAFASDPHIIKAIKONINSVYRDSN 828  
 Db 1084 AELLNGMD--EELDSPELGTRIKRILYK--DEPQYFA-----VYSRIQESQIQPEEG 1135  
 QY 829 INGHDIYKNTVHLVPLNSRGLSNARKFLY--TSATNGDITCSDTTGTQ-----EDTDRY 882  
 Db 1136 I-----LSSTVPIVQASFPFGALTKRIRVGLQAPVEETVKKI 1175  
 QY 883 SONSIGEMTKLQSKTALNRDYY--RRROQRIITQOMFEIDLKLL-----VTEDMSDISP 937  
 Db 1176 -----LGKKAIFSPVLYVEPRRKFHKFTI--MTIPVPPSGEGSVNGYGDATP 1223  
 QY 938 QTRRLINIVSYTG-----RLRANOITFNMORLASIMLTDQWPIRTSML-----I 984  
 Db 1224 NL--RLL--CSITGTSAPQWEDITGTPPLTFIKDCVSTTNVSARF-----WLADCHQV 1274  
 QY 985 LYLEETBGLPDQNTLKTW--YERISKNIPTTKDVEPLLELDGDIRNF----- 1029  
 Db 1275 L--ETVGLASQLYRELICVPYMAKFNVAKNTDP---VESSILRCFCMTDHDVKTLEQ 1328  
 QY 1030 ---EYFLSRFVFLVARDVKTFLPCTVNLDEKLR--EIIDVYARAE-----QINIGG 1077  
 Db 1329 ENEEYARSKDIEVELEKPI--YVDCYGLAPLTKGGQOLVNVNYSFKENRLPFSIKIRD 1386  
 QY 1078 LAYPLP---LHEGPPRPPSGYOPASVYSSASFNGPPGCVSPHSSVYSGLSGPO 1133  
 Db 1387 TSQEPGRLSFLKE--PKTKGLPQTA--VCN-----LNTITPLA 1421  
 QY 1134 HPEYNAANVPATGSSILL-----SMYVDVYCEK--LRQIGLDQNNMPOYCTTTP 1182  
 Db 1422 HKKAERADRQSPASLALRKRYSYLTPESMSPQSCERTIDIMATVAD--HIGLSWTELAR 1480  
 QY 1183 KANINGRVLSOONIDELKKEAMNF--GDM-----HLFSWVLEMRSVSSQVVP 1229  
 Db 1481 ELNFSYDEINQIRVENPNSLISQSPMLLKKWTRRGKNAATTDALTSVLTIKRINDIYVTL 1540  
 QY 1230 EDP-----RFLNENS--SAPVPHGESARRSSHTLPLTELISQTFYTLNFSPEE 1276

Db 1541 EGIIPDYGNISGTRSRADENNHYDV-----DGH--PSFYQLELTPMGLYWT--P 1587  
 QY 1277 LNTLGLDEGAPRHSNLSWOSQTRTPSLSS-----LNSODSIEIS 1317  
 Db 1588 PNFQODD---HFSDISSIESPRTPSRLSDGLVPSQGNLEHPTGPPVTAEDTSLDS 1644  
 QY 1318 KLTDYV---QAEYRQAYREYIAQMSOLEGGTSSITSG-----RSSP--HSTYYIGOS 1365  
 Db 1645 KMDSVTVYDPAPLVDLSDQLKDCQSCQACQASVPGIPNDGRQAEPLRPQTRKVGMS 1704  
 QY 1366 SSGSGHSTLEQERKEGELQKQEDGRKSFIMKRGVDYSSSGVSTNEASPLDPTDEDE 1425  
 Db 1705 SE-----QDEKGSNG-----PDEYVEDKV 1724  
 QY 1426 KSDQSGSKLLPGKSSERPSPLEFQTDILKLGGLRYQLPDSDEDESGTRGVQITPHCSKMI 1485  
 Db 1725 K-----SLFE-DIOLE--EVEAEEMTEDQOQAMLNRYO-----RAEL 1758  
 QY 1486 RTRKRAKQRECAQPOHSAEPTRITKAKLEYLSDDLLOKKSSDSGVRSN----- 1536  
 Db 1759 AMSLQAGWQNEPSPGSLSPQAQARR-----LTGGLDLRDDSDQARDSTLYTGEPE 1811  
 QY 1537 ---ESSPNHSLHNEADDSQLEKANLIELEDEGSHGKRGV-----PHSLSGIQDPTIA 1586  
 Db 1812 GKIEANGNHT-----AEVIEPAKAKPYFPESQNDIGKQSTIKEMLKPKTHCGRTPEPV-- 1864  
 QY 1587 RMCISGDEKKSPESEC--LIASSPESHPACQKAYINLNPSTVTLNNNTAPINRANQNF 1644  
 Db 1865 --SPLTAYQKSLSETSKLVEDAPKPCVPVGMKRM-----TRTTADGKARLNL 1910  
 QY 1645 DEIEG 1649  
 Db 1911 QDEEG 1915

# RESULT 11

S37431  
 N:Ankyrin 2, neuronal long splice form - human  
 N:Alternate names: ankyrin B, 440K splice form; ankyrin-B, brain ankyrin; non-erythroid  
 N:Contains: ankyrin 2, short form  
 C:Species: Homo sapiens (man)  
 C:Date: 06-Jan-1995 #sequence, revision 06-Jan-1995 #text, change 13-Aug-1999  
 C:Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569  
 R:Chan, W.  
 Submitted to the EMBL Data Library, September 1993  
 A:Reference number: S37431  
 A:Accession: S37431  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-3924 <CH>  
 A:Cross-references: EMBL:Z26634; NID:9406287; PIDN:CAA81387.1; PID:9406288  
 R:Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.  
 J. Cell Biol. 114, 241-253, 1991  
 A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a  
 A:Reference number: A39643; MUID:91302466; PMID:1830053  
 A:Accession: A39643  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2077 <OT1>  
 A:Cross-references: GB:X56957  
 A:Accession: B39643  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1443, 3585-3924 <OT>  
 A:Cross-references: EMBL:X56958  
 R:Se, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,  
 Genomics 10, 858-866, 1991  
 A:Title: Isolation and chromosomal localization of a novel moneythroid ankyrin gene.  
 A:Reference number: A40334; MUID:92009921; PMID:1833308  
 A:Accession: A40334  
 A:Molecule type: DNA  
 A:Residues: 463-474, 'PE', 477-495 <TSE>

A:Cross-references: GB:M37123; NID:9178647; PIDN:AAA62828.1; PID:9178648  
 R:Chan, W.; Kordeli, E.; Bennett, V.  
 J. Cell Biol. 123, 1463-1473, 1993  
 A:Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and  
 A:Reference number: A49462; MUID:94075409; PMID:8253844  
 A:Accession: A49462  
 A:Status: preliminary  
 A:Molecule type: nucleic acid sequence not shown  
 A:Residues: 1-3924 <RES>  
 A:Cross-references: EMBL:Z26634; NID:9406287; PIDN:CAA81387.1; PID:9406288  
 C:Genetics:  
 A:Gene: GDB:ANK2  
 A:Cross-references: GDB:127607; OMIM:106410  
 A:Map position: 4q25-4q27  
 C:Superfamily: ankyrin, ankyrin repeat homology  
 C:Keywords: alternative splicing  
 F:2-3924/Product: ankyrin 2, long form #status predicted <MAT>  
 F:2-1443, 3585-3924/Product: ankyrin 2, short form #status predicted <MA2>  
 F:63-95/Domain: ankyrin repeat homology <AN01>  
 F:96-128/Domain: ankyrin repeat homology <AN02>  
 F:129-161/Domain: ankyrin repeat homology <AN03>  
 F:162-190/Domain: ankyrin repeat homology <AN04>  
 F:191-223/Domain: ankyrin repeat homology <AN05>  
 F:232-264/Domain: ankyrin repeat homology <AN06>  
 F:265-297/Domain: ankyrin repeat homology <AN07>  
 F:298-330/Domain: ankyrin repeat homology <AN08>  
 F:331-363/Domain: ankyrin repeat homology <AN09>  
 F:364-396/Domain: ankyrin repeat homology <AN10>  
 F:397-429/Domain: ankyrin repeat homology <AN11>  
 F:430-462/Domain: ankyrin repeat homology <AN12>  
 F:463-495/Domain: ankyrin repeat homology <AN13>  
 F:496-528/Domain: ankyrin repeat homology <AN14>  
 F:529-561/Domain: ankyrin repeat homology <AN15>  
 F:562-594/Domain: ankyrin repeat homology <AN16>  
 F:595-627/Domain: ankyrin repeat homology <AN17>  
 F:628-660/Domain: ankyrin repeat homology <AN18>  
 F:661-693/Domain: ankyrin repeat homology <AN19>  
 F:694-726/Domain: ankyrin repeat homology <AN20>  
 F:727-759/Domain: ankyrin repeat homology <AN21>  
 F:760-792/Domain: ankyrin repeat homology <AN22>  
 F:793-825/Domain: ankyrin repeat homology <AN23>

Query Match  
 Best Local Similarity 20.4%; Score 550; DB 2; Length 3924;  
 Matches 417; Conservative 279; Mismatches 727; Indels 626; Gaps 78;

QY 14 EEENTPALALEKCDVDVERNECGTPTPLMAEQGNVIVELLKNGANCULEDDMT 73  
 Db 73 KEHVGELVOELLGRSSVDSATRKGNLTALHSLAGQAEVVLVKEGANINAQSONGPT 132  
 QY 74 ALISASKEGHIHVEELKSGAS-----LEHRDMG--- 103  
 Db 133 PLYMAAENHIDVYKILLENGANQSTATEDGTPPLAVALAQQHNGQAVALLLENDKGRV 192  
 QY 104 -----GWTALMMACTGRTDVELL 123  
 Db 193 LPALHIAARKDPTKSAALLLQNDHNDADVOSKMMVNTTESGTPPLHIAHYGNVAVATLL 252  
 QY 124 LSHGAMPSTYGLQYSTYPTIIMAGRGHADIYHLLQNGAKVCSXKGTPTPLVMAARG- 182  
 Db 253 LNRGAADVETA-RNGITPLHVASKRGNTVMVKLLDRGGQIDAKTRDGLTPHCAARSQH 311  
 QY 183 -----HLECVKHLAMGADVDEGANSMTALIV 210  
 Db 312 DOYVELLEERGAPLARTKNGSLPLMAAQGHVECVKLLDHKAPYDVDTIDYLFALHY 371  
 QY 211 AVKGTQSYKELKRNPNVNLTKDGNALMAISKE----- 247  
 Db 372 AACHGVRYTKLLDRANPNALNGFPLHACKKRIKIMELLYKXGASIQATSESG 431  
 QY 248 -----GHETVODLLDAGTYVINIDRSGDVLIGAVGSHVEIVRALLQKXADID 297  
 Db 432 LPPIHVAAPMGHLNIVLLLLQNGASPDVYINIGETALHMAARAGOVEVRCILRGALVD 491

QY 298 IRGDNKALYMAVEKGNATWYRLQCNPDTEICTKGEPLIKATKRNIEVELLD 357  
 Db 492 AARREQPTLHIASRLGKTEIVOLLQNHADPAATNGYPLHISAREGOVDASVILE 551  
 QY 358 KCAKVSAYDKKDDTPLHVAIRGRSRLAELLR-----NPKD 394  
 Db 552 AGAANSIATKKGFTPLHVAKAGISLDVAKLLQRAADASAGKNLTPLHVAHNDOKV 611  
 QY 395 GILL-----YRPNKAGETPNIDC-SHOKSILTQI--FGAR-----HLSPT 432  
 Db 612 ALLLEKASPHATKNGTTPHLIAKKNOMQIASLTLYNGAETNIVTKOGVPLHLASQ 671  
 QY 433 ETDGWL-----GYDYSALADLSEPTWQPCVGLYAMWGSKFLKLEDEK-- 484  
 Db 672 EGHDTMVTLLDKGNIMHSTKSLTS-----LHLAAQ-----EDKVN 710  
 QY 485 -----KTPAGQTEPLFOFSWLIVFLTLCC--GGIGLV-FAFVDINLAISISPLAL 536  
 Db 711 ADILTRKGADDAHTRKLGTYPLIV-----ACHYGVKVNFLKGGANVNAKTKNGTYPL 765  
 QY 537 -----IYIFYIYFGRGESGMWMAALSTRLARHIGYLEFLKLFVNP-- 584  
 Db 766 HOAAOQHTHINVLQGAKPNAATANGNTALA--IAKRLGYSIVDVLKVVTEVYTT 823  
 QY 585 -----ELPEOTRALPYRFLFTDYNRLSVGGETSLAEMATLSDACEREFGLA 634  
 Db 824 TTTTEKHLNVPETMEVLDVDSDEGD-DIMTGCGGYLRLPELDELKELGDSLSRSPDLD 882  
 QY 635 TRLEFVFTRESQKKKKKCTCCLPSEVIFLEIVGCIAGITLLAFRVDPRKHLTVAIL 694  
 Db 883 GNNYLRYSLEGRSDS-----LTSFSDSHSLTSKASY 915  
 QY 695 ISIASVGLAFV-----NCRFTW--QYLDLSLQKRRLHSAASKHL 737  
 Db 916 LKDSAVMDSVIYBHOVSTLAKERNSRYLSWGTENLDVAAS--SSPH-- 965  
 QY 738 KSEGMVYKCEVELMAEMAKTIDFTONQRLVYIIGDLACEQDKVLOMLDTRVYLS 797  
 Db 966 --SGFLVIFMVDAR-----GGAMRCRHNGLR--IIPRCKTAPTRVTCGLVVRHLAT 1016  
 QY 798 KGFPIAIFASDPHIIRKAINONLSVLDNSINGHDYRN--IYHLPVFLNSRGISNAK 856  
 Db 1017 MPWVE-----GGLASRLLEVSPSGAQFLGPYIVIPHEPALRG--KEREL 1061  
 QY 857 LVTSATNGDI-----TCSDTGTQEDTDRVYSONSLGEMTKGSKTALNRDTPRRROMOR 912  
 Db 1062 VYLRSENGDSWKHEHCOT-----EDELNELLN--GMDEVLDSPDELEKRIK 1108  
 QY 913 TTRQMSFDLTKLVTEWESDISPQMRLLINIVSYTGRLLRN-----QITENMDRLA 967  
 Db 1109 IITRFE-----POYFAVVS--RIKODSNLIGEGVALSTVPOVQAVPEGALT 1156  
 QY 968 SWINL--TEQWPRYSMTLILETEEGLDQMTL-----KTMVERSKINPTTK--DYEPL 1019  
 Db 1157 KRIRVGLQAPMHSSEIVAKKILGNKATSPVLTLEPRRKRKFKPLTMTTPPKVASSD--ML 1215  
 QY 1020 LEIDGDIRNEFV-----LSSRTPVLAADVKTPLCTVNLDPKLETIAD 1065  
 Db 1216 NGFGDAPTLRLLCISITGTPPAQMEDITGTPPLFVNECSF--TTNV--SAREPLID 1270  
 QY 1066 VRAAREQINIGGLAVPLPLHEGPPRPSGYSQASVSSASFGNGPPGGVSPQSHSY 1125  
 Db 1271 CRQIOESVTFASOYVREL-----IC-----VPYMAK 1296  
 QY 1126 YSGSGPOHPFYNNRAAVPATGSSLLSMYDVYCEKRLQEGLDQMMPOYCTTIKAN 1185  
 Db 1297 FVYRAKSHDPLEAR-----LRCKMDDKDKTLEQ-----QENFEVARSQVEV 1342  
 QY 1186 INGR-VLSQC--NIDELKEKAMNFGDMLFRSVATLMSRVSQVVEDEPRLTENSAP 1242  
 Db 1343 LEGKPIYVDCGNLVPLTKS-----GQHIFSFPAFK-----ENNL--PLFYKVRDTQ 1389

QY 1243 VPHESSRRSSHTEPLTELSQTPYTLNFSEELNTLGLDECAPRHSNI----- 1292  
 Db 1390 EPCGR-----LSFMKEPKSTRGLVHOAICNLTITLTYKESR 1427  
 QY 1293 SWSQOTRTPSLSSINSQSSSIEISKRLDKVQVQAEYND--AREKIAQMSQLEGGSSTIS 1351  
 Db 1428 SDQEQEELIMTSKNDDETSTETSYLKSILVNEVPVYLAAPDLSEVSEKQDILMT-- 1485  
 QY 1352 GRSSPHSTYYIGOSSGSIH-----STLEOERKGEKELKQEDGRKSFMLKRGVDIYSS 1406  
 Db 1486 -----ALLTPVDSKAGSIKVEKELYKAABEERGEPEIYER--VKEDEKVEIL--R 1534  
 QY 1407 SGVSTNASPLDP-----ITEE-----DEKDSQSKL-----LPG 1437  
 Db 1535 SGCTRDRESSVOSSRSEGLVEEEMWIVSDIEIEEROKAPLEITEYPCVEVRIDKEING 1594  
 QY 1438 KKSERPSL--FOTD-----LTKKG--GLRYOK-----LPSDD 1468  
 Db 1595 KVEKDSGLVNYLTDLNTCVPLPKRQLQTVQDQAKKCEALAVGSSSEKGDIPDET 1654  
 QY 1469 ESGTGRVQITPHCSKMRIRLAKAKORECAPSPESHSAEPIRPFIRAKLEYLSDALDKDS 1528  
 Db 1655 QSTOKHKKPSTGIKKPVR--KTLKEKOKKEEGLOASAE-----KAEL-----KGS 1699  
 QY 1529 SDGVRNENESSPHNSLHNEAADSOLE-----KANLELDEBHGSKRGMPHL 1577  
 Db 1700 SEESIGDPLAPLEPLPYKATSPLEPPIGSIKRYALOKRVED--QKGRKIPRV 1758  
 QY 1578 SGLDPLIARMSICSEDKSPSECSLIASSPESBPACAKAYNLNTPSTVTLNNNTAPT 1637  
 Db 1759 KGED-----VPKKTIRHPAAS-----PSLSERAPSPSPKTERHSTLSS 1802  
 QY 1638 NРАНQNFDELIGIRISQVILRGPSP--NPTAVQENLKSMAKRSQBSYRLSKDASE 1696  
 Db 1803 SATERHPVPSSTKTEK-----HSPVSPSA-----KTERHSPASSSKT-----E 1843  
 QY 1697 LHAASSEST 1705  
 Db 1844 KHSVPSPST 1852  
 RESULT 12  
 149502  
 ankyrin - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 13-Aug-1999  
 C:Accession: 149502  
 R:White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.  
 Mamm. Genome 3, 281-285, 1992  
 A:Title: Murine erythrocyte ankyrin CDNA: Highly conserved regions of the regulatory  
 A:Reference number: 149502; M0ID:92345717; PMID:1386265  
 A:Accession: 149502  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1862 <RES>  
 A:Cross-references: GB:M84756; NID:9191939; PIDN:AAA37236.1; PTD:9191940  
 C:Gene: Ank-1  
 C:Superfamily: ankyrin; ankyrin repeat homology  
 C:Keywords: alternative splicing  
 F:40-72/Domain: ankyrin repeat homology <AN01>  
 F:73-105/Domain: ankyrin repeat homology <AN02>  
 F:106-138/Domain: ankyrin repeat homology <AN03>  
 F:139-167/Domain: ankyrin repeat homology <AN04>  
 F:168-200/Domain: ankyrin repeat homology <AN05>  
 F:201-233/Domain: ankyrin repeat homology <AN06>  
 F:234-266/Domain: ankyrin repeat homology <AN07>  
 F:267-299/Domain: ankyrin repeat homology <AN08>  
 F:300-332/Domain: ankyrin repeat homology <AN09>  
 F:333-365/Domain: ankyrin repeat homology <AN10>  
 F:366-398/Domain: ankyrin repeat homology <AN11>  
 F:399-431/Domain: ankyrin repeat homology <AN12>  
 F:432-464/Domain: ankyrin repeat homology <AN13>



1648 -EGIRETSQVILRPGSPNPTA-----YQENNLKSMARKRSQSSYT 1688  
1751 QKQPPETGSP--KAGEPRLMAPESAFSOEVQDELQNPGEQVTEQFT 1798

RESULT 13  
A55575  
Ankyrin 3, long splice form - human  
N:Alternate names: ankyrin G  
C:Species: Homo sapiens (man)  
C:Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 20-Sep-1999  
C:Accession: A55575  
R:Kordeli, E.; Lambert, S.; Bennett, V.  
J. Biol. Chem. 270, 2352-2359, 1995  
A:Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the ax  
A:Reference number: A55575; M0ID:95138209; PMID:7836469  
A:Accession: A55575  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-4377 <KOR>  
A:Cross-references: GB:U13616; NID:9608024; PIDN:AAA64834.1; PID:9608025  
Genetics:  
A:Gene: GDB:ANK3  
A:Cross-references: GDB:424503; OMIM:600465  
A:Map position: 10q21-10q21  
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology  
C:Keywords: alternative splicing; peripheral membrane protein  
F:106-138/Domain: ankyrin repeat homology <AN01>  
F:139-171/Domain: ankyrin repeat homology <AN02>  
F:172-200/Domain: ankyrin repeat homology <AN03>  
F:201-233/Domain: ankyrin repeat homology <AN04>  
F:234-266/Domain: ankyrin repeat homology <AN05>  
F:267-299/Domain: ankyrin repeat homology <AN06>  
F:300-332/Domain: ankyrin repeat homology <AN07>  
F:333-365/Domain: ankyrin repeat homology <AN08>  
F:366-398/Domain: ankyrin repeat homology <AN09>  
F:399-431/Domain: ankyrin repeat homology <AN10>  
F:432-464/Domain: ankyrin repeat homology <AN11>  
F:465-497/Domain: ankyrin repeat homology <AN12>  
F:498-530/Domain: ankyrin repeat homology <AN13>  
F:531-563/Domain: ankyrin repeat homology <AN14>  
F:564-596/Domain: ankyrin repeat homology <AN15>  
F:597-629/Domain: ankyrin repeat homology <AN16>  
F:630-662/Domain: ankyrin repeat homology <AN17>  
F:663-695/Domain: ankyrin repeat homology <AN18>  
F:696-728/Domain: ankyrin repeat homology <AN19>  
F:729-761/Domain: ankyrin repeat homology <AN20>  
F:762-794/Domain: ankyrin repeat homology <AN21>  
F:795-827/Domain: ankyrin repeat homology <AN22>  
F:828-860/Domain: ankyrin repeat homology <AN23>

Query Match 5.9%; Score 526; DB 2; Length 4377;  
Best Local Similarity 20.0%; Pred. No. 9.1e-22;  
Matches 416; Conservative 266; Mismatches 664; Indels 734; Gaps 80;

QY 14 EENIPALKALLCKDVERNECGOTPLMLAEOGWEIYKELIKNGANCLEDLNMT 73  
DB 83 KEHVEVVELLQREANVDATKKGNTALHIALAGAEVVKVLTVGANVNAQSQGFT 142  
QY 74 ALISKRGHIIYBELLSGASLEHMDGWTALMNAKCGRTDVELL----- 124  
DB 143 PLVMAAOENHLEVVKFLDNGASOSLATEDEFTPLAVALOQGHQVVSLLENDTKGVR 202  
QY 125 -----SHGANPSVTG----- 135  
DB 203 LPAHIAARKDTRAAALLLQNNADVESGFTPLIAHAGINIVATLLNRAAVD 262  
QY 136 ---OYVYPIIWAAGRGHADVHLLQNGAKVNCSDKYGTTPVWMAKRG----- 182  
DB 263 FTARNDTPLHVAASKRGANVVKLLDRGAKIDAKTGDGLPLHCGARSGHEQVEMLLD 322  
QY 183 -----HLECVKHLAMGADVDOEGANSMTALIVAVKGYTOS 219

DB 323 RAAPILSKTYKNGSLPLHMAIQDGLNCVQLLQHNVPVDVNDYTLALHVAHGHYKV 382  
QY 220 VKEILKRNPNVNLTDKQNTALMASKE----- 247  
DB 383 AKVLLDKKAPNMAKALGFTPLHACKKNIKYMEILLKKGASIQAVTESGLPIHVAAP 442  
QY 248 -GHTEIVQDLLDACTYVNPDRSGDYVLGAVRGHVEIVRALLQKADIDRGQDNKA 306  
DB 443 MGHVNIYSQMLHGHASPDNTNVGRTALHMAARSGAEVRYLYODGAEVAAKADQPR 502  
QY 307 LYMAVEKGNTPVYDILQ--CNPD-----EIGM 333  
DB 503 LHISARLGRKADIVQQLLOQASRNAATTSQYTLHLSAREGHEDVAFLDHCASLSTT 562  
QY 334 KDGFTPLIKATKKNIEVVELLDKGAQVAVDKDPTPLHVAIRGRSRLLELLRNPK 393  
DB 563 KKGFTPLHVAAKKGKLEVANILLQKASPDAAKSGSLTPLHVAHVDNQKVALLL--D 619  
QY 394 DGRLLYPRNKAQETPRYVND--SHQSLIQI--FGAR-----HLSFTEDGM 438  
DB 620 QGASPHAAAKNGYTPPLHIAKKQMDIATTLLEYADANAVTRQGIASVHLAAOECHVDM 679  
QY 439 LGYDLVSSALADILSEPTQ-----PICVGLY 466  
DB 680 VSLILGRNAVNLNKSGLTPLHIAQEDRVNAEVLVNOGAHVDAQTKMGTPLHVGCH 739  
QY 467 AQWGSCK--SFLIK-KLEDEMKTPAQOQEPLEFQFSWL--YVFLLLCGGLGVFAP- 520  
DB 740 --YGNIKIVFLLHSHAKVNAKTKNGY--TPLQAQOQGHTHINLVNN--ASP 790  
QY 521 ---VDPLAIAISLFLALYIFVIYVYGGREGSGNMMAVLSRLRNHGYLELTK 577  
DB 791 ELVYNGTALGT-----AKRGYISVDGT 814  
QY 578 LMFV-----NPE-----LPEQTKALPVAFLETDNRLSV--GGET 613  
DB 815 LKIVTEETMTTIVTEKKNVPEYTNNEVLDMSDDEVRAKANAEMLSDEYISDVEGSD 874  
QY 614 SLA-----EMATLSACEREGFLATRL-----FRVFRYESGKKKKKTKCCL 658  
DB 875 AMGDGDKYLGPDLELGDLSLPAGYMGFSLGARSASLRSSDSRSTYLN--SSYAR 932  
QY 659 PSFVIFLEFYGCITAGITTLAIFRVPD--KHLT-----VNAIISIVAGLAFVLC 709  
DB 933 DSMIBELVPSKQDILFTFRSDSLNHYSAADTLNVLNVSSPIHSGLVSPMDA 992  
QY 710 RTWQVYLDLSLNSQRRLHSAASKLHLKSEGFMYL--KCEVELMARMAKTIDSTFO 765  
DB 993 RG-----GSMRGSRHNG-----MRTIIPRKTAP-----TR 1019  
QY 766 NOTRLVYIIDGLDACQDQVLDMLDVRVLFSGKPFIALFASDPHIIKAINQNLNVSR 825  
DB 1020 ITCRLV-----KRHLAN-----PPH-----GERRGISRSLV 1047  
QY 826 DSNINCHDYMRN--IVLPEVLSRGLSNARKEFLVTSATNGDITCSQDTQCDTPRVSQ 884  
DB 1048 EMGPRAQGLPRVIVIEIPFGSMRG--KRELIVLRSENG-----TKKEQFD---SK 1096  
QY 885 NSLGEYTKL--GSKTALNRDYYRRRQORTTROMSFDTKLVTEDWFSIDSPQMYR 942  
DB 1097 NE--DLTELLNGDELDSPFEIGKRIRITKDF-----POYFAVVS--RIKQ 1142  
QY 943 LNIIVSVTGRILRAN-----QITFMNDRLASVINLTDQMPYTSMLIYLETBGLPQM 997  
DB 1143 ESNQIGPEGLISSTTVPLVQASFEQALTKRIRVGLQ-----AQVYPEI 1188  
QY 998 TLKTM-----YERISKNP-----TKDVPLEI 1022  
DB 1189 VKKILGNKATPSPIYVBERRRKKFKRPIITMTIPVPPSGEGVSNKYKQDTPNLLCSI 1248  
QY 1023 DG-----DIRNEFVLSRFPVLVADVTFPLPCTVNDLPKRLRIADVRAAREQIN 1074

Db 1249 TGGTSPAQWEDI-----TGTPPLTFIKDCVSF---TTNV--SARFWLADCHQVLETVG 1296  
 Qy 1075 IGLAVPPLPLHGGPPRPSPGSGQSPASVCSASFNGPFGGVSPQPHSSVYSGLSGPGH 1134  
 Db 1297 LAFQYREL-----ICV-----PYMAKEY-----VFAMKMDP-- 1323  
 Qy 1135 PFYRAVAPATGSSLLSMTVDVCEKLRQIEGLDONMMPQYCTTIKKANINGRVLSDQC 1194  
 Db 1324 -----VESLRNFCMTDVKDKLQOE----- 1346  
 Qy 1195 NIDELKEMAMNGDWHLFRSMVLEMRVSQVVPEDPFLN-ENSSAPVHGE-----S 1248  
 Db 1347 NFEEV-----ARSKDIEVLEKGPVYDVGYNLAPLTKGQOOLVFN 1386  
 Qy 1249 ARRSHTLP-----LPELSOTPYTLNFSEELNTGLDGAHRHSLMSQSTRRPPL 1304  
 Db 1387 FYSKENRKLPPSIKIRTSQPCGRUSFLKERTTKGLPOTACNMLITLPAKKETES- 1445  
 Qy 1305 SSLSNDSISIEISKLDKQVQAEYDAYRE---YIAQMSOLEGGTSGSTISGRSPHSTVY 1361  
 Db 1446 ----DODEIE---KTDRQSPFASIALRKRYSTLTERGMIERGTAT---RSLP---TTY 1492  
 Qy 1362 ICGSSSGGSIHSTLEQERKGEKQEDGRKSFMLKRGVYIDYSSGVSINERAPLDPIT 1421  
 Db 1493 ----SYKPFESTRPQSWTAPITVPGPAKS-----GPTSLSSSSSNTSPASPLKSIW 1541  
 Qy 1422 EDEKSDQSGSKILLPGKSSRPSPLOTDLKLGGLRQKLPDDESGTGRVQIPLPHC 1481  
 Db 1542 -----SYTSPPIKSTL-----GASTTSSVKSISDV 1567  
 Qy 1482 SKMIRTRRLAKQRECAPQESHSAPIRTFIKAKEY---LSDALDK-----KDS 1528  
 Db 1568 ASPIRSLRT-----MSSPIKTVVQSOPVNIQVSSGLANAPAVTEATPLKGL 1614  
 Qy 1529 SDGVRNNESSPNHSLNEADDSQLEKANILEDEHGCKGKMPHS----- 1576  
 Db 1615 ASNSTFSSRTSP-----VTTAGSLERSSITMTTPAPSPKNIMYSSSLPFKSIITSA 1668  
 Qy 1577 -----LSGLDDPPIAHMSICSEDKSPSECSLIASPPESMWA----- 1614  
 Db 1669 PLISPLKSVSPKSVKDYVSISSAKITMASS---LSSPVKOMRHAVALVNGSISPLKY 1725  
 Qy 1615 CQKAYNLRTPSTVTLNNN-TAPTRANO-----NFEDEIGIRETSOYI---LPPGSPN 1665  
 Db 1726 ASSSTLLNGCKATATLEKISSATNSVSVSAATDIVEKVFSTTAMPSPSLKSYSA 1785  
 Qy 1666 PTAQONENKMAHKBORSSTYRLSDASFLHAASEST 1705  
 Db 1786 PSAFOSLRTPSAS-----ALYTSL---GSSISATTSVYT 1816

## RESULT 14

S37771

ankyrin, erythrocyte - mouse

C:Species: Mus musculus (house mouse)

C:Date: 27-May-1994 #sequence\_revision 03-Aug-1995 #text\_change 13-Aug-1999

C:Accession: S37771

R:Bikemeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.

J. Biol. Chem. 268, 9533-9540, 1993

A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found an

A:Reference number: S37771; MUID:93252825; PMID:8486643

A:Accession: S37771

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1848 &lt;BIR&gt;

A:Cross-references: EMBL:X69063; NID:9311816; PIDDN:CAA4801.1; PID:9311817

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

F:48-80/Domain: ankyrin repeat homology &lt;AN01&gt;

F:81-113/Domain: ankyrin repeat homology &lt;AN02&gt;

F:114-146/Domain: ankyrin repeat homology &lt;AN03&gt;

F:147-175/Domain: ankyrin repeat homology &lt;AN04&gt;

F:176-208/Domain: ankyrin repeat homology &lt;AN05&gt;

F:209-241/Domain: ankyrin repeat homology <AN06>  
 F:242-274/Domain: ankyrin repeat homology <AN07>  
 F:275-307/Domain: ankyrin repeat homology <AN08>  
 F:308-340/Domain: ankyrin repeat homology <AN09>  
 F:341-373/Domain: ankyrin repeat homology <AN10>  
 F:374-406/Domain: ankyrin repeat homology <AN11>  
 F:407-439/Domain: ankyrin repeat homology <AN12>  
 F:440-472/Domain: ankyrin repeat homology <AN13>  
 F:473-505/Domain: ankyrin repeat homology <AN14>  
 F:506-538/Domain: ankyrin repeat homology <AN15>  
 F:539-571/Domain: ankyrin repeat homology <AN16>  
 F:572-604/Domain: ankyrin repeat homology <AN17>  
 F:605-637/Domain: ankyrin repeat homology <AN18>  
 F:638-670/Domain: ankyrin repeat homology <AN19>  
 F:671-703/Domain: ankyrin repeat homology <AN20>  
 F:704-736/Domain: ankyrin repeat homology <AN21>  
 F:737-769/Domain: ankyrin repeat homology <AN22>  
 F:770-802/Domain: ankyrin repeat homology <AN23>

## Query Match

Best Local Similarity 19.6%; Pred. No. 7.8e-22;

Matches 397; Conservative 281; Mismatches 651; Indels 697; Gaps 81;

Qy 21 LKALLECKDVERNECCGTPPLMAEGGVEIKELKNGANCULEDDNTALISASK 80  
 Db 98 VRELIVYGANVNAOSQKGFPLMAQENHLEVFLENGANQVATEDGETPLAVALQ 157  
 Qy 81 ECH----- 83  
 Db 158 QGHENVVAHLINYGTKGKRLPALIAANDTRPAANVLLDPPNDVLSKGFPLPIA 217  
 Qy 84 ----THIVEELKSGASEHRDMGWTALMAACYGRDVELLSHGANSVTGLQYS 138  
 Db 218 AHYENINAVQALLNBRASVNPFPQNGITPLLIASRGVNIYVRLLDRAQOE--TRKDE 276  
 Qy 139 VPIIAGRGADIVHLLONGAKYCSDKYGTTPVLAARKGHLECYKHLIAGADVD 198  
 Db 277 LPLHCAARGHVRISIEILLDHGAPIOAKTKGLSPIHMAAGDHDICVRLLOVNAELD 336  
 Qy 199 -----QEGA-----NSMTALIVAKGYQSYKEILK 225  
 Db 337 DITLDHLPLHAHAGCHHRVAKVLLDKAKRKSRLNFTPLHACKNHLRWELLIK 396  
 Qy 226 RNPVNLTDKDGWTALMAISKEGHEIVODLDAGTYVNIIPRSGDVLIGAARGHVEI 285  
 Db 397 TGAIDAVTESGLTPLHVASFMGHLPIYKLLDGRGASPNVSVKVTPLHMAARAGHVEY 456  
 Qy 286 VRALLQYADIDIRGDNKTALYAVEKGNATVRIIDQNPDTLCTDGETPLIKATK 345  
 Db 457 AKYLLQKAKANAKADDDPLHCAARIGHGTGVKLLDNGASPNLATTAGHTPLTAAR 516  
 Qy 346 MRNIEVEVLELLDGKAVSAVDKGDTPPLVAIIGRSRIELLLRPKGRILYPRNAG 405  
 Db 517 EGHVTRALLALEEASQACMTKKGTPLHVAKYGVRLABELLEHDA-----HPNAG 570  
 Qy 406 E--TPYNDICSHQ-----KSLI-----TQIFGAR----- 427  
 Db 571 KNGITPLHVAHNNNDIYKLLPRGSGSPHSPAMNGYTPLHIAKQNOIEVARSLQYOG 630  
 Qy 428 -----HLSPTERDGMIGDYLSMALDIIEPRMOP-----PIC-- 462  
 Db 631 SANAESVGYTPPLHAAQEGHTEMVALLSKQANGMLGNKSGTLPLHLYVSGGHVPAVDY 690  
 Qy 463 -----VGLYQWQSGK--SPLIKLKD--EKKTFAGOOTPELQFSWL 501  
 Db 691 LIKHGTVVADATTMGCTPLHVAHYGNIKIKYVFLDHOADVAKTKLGG--SPLHQAAQ 748  
 Qy 502 --IVFTLLLLGGGLGVAFAPVD-----TNIAISLSPLALITYFFIYIFGGRGE 553  
 Db 749 GHTDIYVTLKNG-----ASPNEVSSNGTTPLAIAKRLYISVTDVIAKYV----- 794  
 Qy 554 SMMWMAALSLRLARHIGYLELEFKLMFVNPPELPBQDTKALPV---RFLFTYINRLS-- 608



```

Db 795 --DERSVYLVDKHMMSYPERVDELIDVS---EDEGTAHISIMGDELIVGSKAEKRSRD 848
QY 609 VGGFSLAMITLSDACEREFGLATLFR-----VFRFE-ESQKKKKKKTKCCL 658
Db 849 VEEKELIDFVFKLDQVVE-----SPAIPRIPCVPTETVIRSEDDQAKSEYDEDESLI 902
QY 659 PEFVLEFVIGCIITGLLAFVDPKHLVFNALISIASVAGLAFALNCRIMQVIVDS 718
Db 903 PS-----SPATEISDNI-SEFVASVHTGFLVSR-----WYDA 933
QY 719 ILNSQRKRLHSAASKLHKLKSEGMKVLCEVELMARMAKTIDSEFTQNGRLVYIIGELD 778
Db 934 RCGSMRSGRHN-----GLRVVIP-----PRICAAPTRTICLVA----- 966
QY 779 ACEQKVLQMLDVTAVLFESKGFALIFASDPHIIITKALNOUNLVLRSDNINGHDMYRN- 837
Db 967 -----KPOKLTTPPL-----ABEELIASIIIALGPTGQOFLSPV 1001
QY 838 IVHLPEVL-NSRGLSNARKFLVTSATNGDITCSPTTGTQEDTDRRVSONSLGEMTKLGSK 896
Db 1002 IVEIDHFASHGSG--DELYVLVLSKNSV-----WKEHKSRYGESYIDQILN-GMD 1049
QY 897 TALNRDYYRRQOMORTTQOMSEDLTKLVLT--EDWFSDISPQTMRLNLNIVSYGRL 953
Db 1050 EELGSLIELEKRRVCRIIT--TDPELYFVMSRLCQD-YDTIGEG-----GSLRSKL 1099
QY 954 LRANOITFNMRLASMINLLEOWPYRISMLILYLEETRGEDDQMTKIMYERISKNIPTT 1013
Db 1100 VPLVQATPEPNAVNTKVKLALQ-----ADPVDELATYKILGNQ----- 1137
QY 1014 KDVEPLLEIDDINFEVFLSSRPVLVARDVKTFELPCTVNLDPKLEIILADVAREQI 1073
Db 1138 -----ATFSP-ITVVEPRRRRFRHPI----- 1157
QY 1074 NIGGLAVPRLPLHBEPRPPRGYSOPAS--VCSASANGPFGCVVSPQSHSYGSLG 1130
Db 1158 --GRILPPLPSWTDNPR-DSGEGDPTSLRLLCVI-----GTTDOAQ-----WEDIT 1201
QY 1131 GPQHPFY-NRAA-----VPATG-----SLLLSMTVQVVCERKLQIGELD 1170
Db 1202 GTTKLIYANECANFTTNVSARFWLSDCPRTAEVHFATLTKELTA----- 1247
QY 1171 QNMMPQYCTTIKKANIN--GRVLSQCNIDE-IKKEMANFGMHLFRSVLTEM----- 1220
Db 1248 --VPYMAKFYIFAKMDAREGRJRCYCMTDKVDKTLQEHENFEVAVARSHDIEVLCMP 1304
QY 1221 --RSVESQVY-----EDPRFLNENSSAPVPGESARR-----S 1252
Db 1305 LFAELSGNLVYVKAQAQORSHFQSFERNRLAIFVKVYDSSREPGEFLSFLRKTKMYEDT 1364
QY 1253 SH-----TELPLETSL-----SOTPYLNFSPFELMNTLGLDEGAPR----- 1288
Db 1365 QHILCHLNTMPCTKSGAGEDRRRTILPLILKYSIISSESLGTSTJTDVEMAMAVIRE 1424
QY 1289 HSNLSWOSQVR-----RTPSLSLNSODSSI-----EISKLTDKVOAEY 1327
Db 1425 HLGSLMAELAELOFSEVEDINRIIVENPNLSLDSTALLTLMTWREGENAKM-ENLYTAL 1483
QY 1328 RDAVREYIAQMSQLEGGTSGSFTLSGRSSPHSTYYIIGSSSGGSHSTLBOERKEGELKO 1387
Db 1484 RNIRSELYVMLE-----VSGROS-----RNLLKER----- 1509
QY 1388 EDGRKSFMLKRGDIVDYS-----SGVSTNEASPLDPIEEDKESQOSSKLLPGKSS 1441
Db 1510 RHGDRREYSLSPSQVNGVSSSLQDELDELSPASLOIYALP-SPLCADQYMWNYTIDALPLAITE 1568
QY 1442 ERPSLFQTDKLGKGGGLRYQLPSDEDESGTGRVQITPHCSKMIKTKRLAKQREGASPO 1501
Db 1569 HDTMLEMSDMQVMSAGLT-PSLYVTAEDSS-----LECSKAE 1603
QY 1502 EHSNEPIRTFIKKEVYISDALDKKDSGVSNSNESSPNHSLHNEAADSOLEKANLIE 1561
Db 1604 D-----SDAIPKWK-----LEGAHSEBTQGPGLGSDOLVE 1633

```

```

QY 1562 LEDEHSGKRGKMPHSLSGLODPITLARKSICSEDKSPSECSLI-----A 1605
Db 1634 DDTVDSDATNGGLADLLG--QORVHARIT-----DSPSVROVLDRSQARTLMDKQSGTA 1685
QY 1606 SSPE-----SW--PACQAKANLNRTPSYTLNNKTPTRANONPDEI-----EGT 1650
Db 1686 VHPQEAQTQSSWQEBVETQGFHSGFORITTTI-----QGEFGALQETQVSVSTREHYQRP 1740
QY 1651 RETSOVLIRPSPNPPTA-----VONENLKSMAHKRSRSSSYT 1688
Db 1741 PETGSP--KAGKEPSLMAPEASAFSQEVQGDDELQNTIGEDVTEQFT 1784

RESULT 15
T42714
ankyrin 3, splice form 2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42714
R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Yialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title: Ankyrins (epithelial ankyrin), a widely distributed new member of the ankyrin ge
the repeat domain.
A:Reference number: 222237; NCID:95340633; PMID:7615634
A:Accession: T42714
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1765 <PEP>
A:Cross-references: EMBL:L40632; NID:9710548; PID:9710551; PIDN:AAB01605.1
A:Experimental source: strain C57BL/6J; kidney
C:Genes:
A:Gene: Ankyrins
A:Map position: 10
A:Introns: 1587/1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 5.8%; Score 510.5; DB 2; Length 1765;
Best Local Similarity 20.1%; Pred. No. 1,6e-21;
Matches 396; Conservative 278; Mismatches 641; Indels 653; Gaps 78;

QY 14 EENINPALKALKECKVDNERECQTPMLAEOGNEIVKELKNGANCULEDDMWT 73
Db 66 KESHEVEVELLQREANVDAARKKGTALHISLNGOAEVAVVLYTNGANVAQSONGFT 125
QY 74 ALISAKKEGHINIVBELKSGASLEHRDNGMTALMACYKRTVVELL----- 124
Db 126 PLYMAOENHLEVREFLDNGASQSLATEDGFTPLVALQGHQDVVSLLENDTKGKVR 185
QY 125 -----SHGANPSYTGK----- 135
Db 186 LPLALTAARKDDTKAALLLNDNTNADVESKSGFTPLHIAHYGININATLLNRAAAD 245
QY 136 ---QYSVYPIIWAAGGHADIVHLLQNGAKVNSDKTGYTPLYVAARKG----- 182
Db 246 FPARNDITPLHNASRGNANMKLLDRKAKIDATRGITPLPHCGASGHEQVYVELMD 305
QY 183 -----HLECYKHLIACADVQDEGANSMTALIVAKGYTQS 219
Db 306 RSAPILSKTKNGLSPLHMATQGDHNCVQLLQHNVPVDVYNDYLTALHYAAGCHYKV 365
QY 220 VKEILKRNPNVLTDKDGTALMISKE----- 247
Db 366 ARVLIDKASPAKALNGFTPLHACKNRIRVMEILLKHGASIOAVTESGLTPHYAAV 425
QY 248 -GHIEVODLADAGTYVNIPIRDSGDTVLIGAVGGHVEITVALLQYADIDITRODQNTA 306
Db 426 MCHVAVIVSGLMHHGASPMPTVNRGFTALHMAARSOAEVAVVLYVODGQVAKAKDQTP 485
QY 307 LTVAAVEGNATVNRDILQCNPDTEICTYKDEGTPLIKATKMRNIEVELLDGKAVSAVD 366
Db 486 LHISARLGRKADIVQOLLQOGASPNNAATTSGYTPLHIAAREGHEVAVFAELDHGASLSTYT 545

```



```

QY 367 KKGDPLHVAIRGRSRLAELLR-----NPKDGRLL----- 398
DB 546 KKGDPLHVAIRGRSRLAELLR-----NPKDGRLL----- 398
QY 399 --VRPNAGETPNIDC-SHOKSILNOI-FGAR-----HLSLETEDGMIGY 441
DB 606 SPHAAKNGYTPPLHIAKKOMDIATSLLEGADANAVTQGLASVHLAQBEGHVDVSL 665
QY 442 LXXSSALADILSEPTMOP-----PICVGLYAMQ 469
DB 666 LLSHNAVNLNKSGLPLHLAAQEDRVNAEVLVNOGAHVDAQTKMGYPLHVGH--Y 723
QY 470 GSGK--SFLK--KLEDEKKTFRAGQOTEPLOFSMLIVFLT-----LLOGGLGV 516
DB 724 GNKIVFPLHSAKVAKTKNGYTAHQAQGHHTHINVLIQNNASPNELTVNG-- 779
QY 517 FAFVDTNLAIATSLFLALITYEIVY----- 544
DB 780 -----NTALAIARLIGYISVVDLKYVTEELMTTTEKHKMNVPTMNEVLMSDEV 834
QY 545 -----YFGRREGESWMMALSTRLAHIGYLELFLK-----MEVN 562
DB 835 RKASAPKRLSDGEYISDGEEDKCTWPKIPKVO-----EVLKSEDAITGDTDYLG 886
QY 583 PPELBEQTTALPVRFLETDYNNRLSSVGETSLAEMITATLSDACEREGFLATPLFVFR 642
DB 887 PDLKEIGDSDLPAR-----GYVG--FSLGARSASLR-----SFSSDRSYTLNR 928
QY 643 TEESQGRKKWKCTCLPSFVIFLFTVGCIIAGITLAIIFRYP-KHLY-----VNAI 693
DB 929 SSYAR-----DSMIEELLPVSKBQHLLFTREPDSLSRHSMADTLDNVNLV 977
QY 694 LISTAVVGLAFVNLCTWMOVLDLSLNSQ-----RK-----RLHSASKLHL- 737
DB 978 SSPVHSGFLVSFVNDARG-----GSMGRSHHGMRIIPRKCCTAPRTICRLVKKHKL 1032
QY 738 -----KSEB-----FMKVLKCEVELMARMAKATIDSFON 766
DB 1033 NPPWVEGEGLASKLVEKMPAGAOFLGPVIVEIHFSGMGRKRELIVLSSENETYKH 1092
QY 767 Q-----TFLVYIDGLDACEODKVLQMLDLYRV--LFSKGFPIAIFASDPHIIKAINON 819
DB 1093 QFDSKNEDIAELLNGMD--EELDSPEELGTKRICRITK-DEPOYFA-----VYSRIKOE 1144
QY 820 LNSVLBDSNINGHYMRNIYHLPVFLNSRGLSNARKFLV--TSATNGDITCSDTGTQ--- 875
DB 1145 SNOIGPEGI-----LSTTVPLVOASPEEGALTKRIRVGLQAOP 1184
QY 876 --EDIDRRVSONSLGEMTKLGSKTALNROTY--RRROMORTITROMSFDLTKL--VT 928
DB 1185 VPEELVKRI-----LGNKATFSPVIVEPRRRKFKHPIT--MTIIPVPPSGEGVS 1232
QY 929 EDWFSDISPQTRRLINIVSYTG-----RLRANOITFNMRLASWINLTDOWPYRT 980
DB 1233 NGYKGDATPNL--RLT--CSITGTSAPQWEDITGTTPLEIKDCVSFTVNSARF--- 1284
QY 981 SWL-----ILYLEFEGLPDOMTLKMTYERISKNIPTTKDVEPLLEIDGDLRNEFVLS 1035
DB 1285 -WLADCHQVL--ETVGLASQLYRELIC-----VPY-----MAKFVFAKT 1321
QY 1036 RTPVLVARDVKTFLPCVTNLPKLEITADYARAREQINIGLA-YPLPLHEGPPRPPS 1094
DB 1322 NDP-----VSSSLKCFGMDTRDKTL-----EQDENFEVARSKDIEVLEBKPIYVD 1369
QY 1095 GYSQAPAVCSSASFNGPFGVSPQPHSYSGLSGFQHPF--YNRAAVPATGSSSL 1150
DB 1370 CYGNALPLTK-----GG--QOLVFNFFS--FKENRLPFPSIKIRDTSOEPCGRLSFL 1416
QY 1151 LSSMTVDVCEKLRQIELDQNMMPQYCTTIKKANINGRVLSQCNIDELKEMAMNFGDW 1210
DB 1417 KEPKTT-----KGLPOTAV-----CNLNTLPRAHKAERAKDRQS----- 1451

```

```

QY 1211 HLFPSVLEMR--SVESQVPEDEPRFLNENSSAPVPGESARSSHTLEPLTSLQTP 1267
DB 1452 --FASIALKKRYSYLETPEMSQSPQPCERDIMALY-----ADHGLSTELARE-- 1499
QY 1268 YTLNFSFEELNLTGDEGAPRHSNLS-----W--QOQTRTPSSLSSLNQDSSTIEI 1316
DB 1500 --LNFVDEINQIRVEN--PNSLISQSPMLKKNVTRDGNKATDALTSVLTAKINRIDI 1554
QY 1317 SKLTDKVOAEYRDATREYIAQMSQLEGGSGSTIGRSSPHSTYIIG--QSSSGSHTST 1374
DB 1555 VTLLEGPIDYGNL-----GTRSFADENNVEFHDVYDGMQNETPSSGLSP 1600
QY 1375 LEOERKGELEKOEORRKSFLMKRGVDIDYSSSGVSTNEASPLDPTTEDEKSDQSG-- 1431
DB 1601 AQARRLTGLDLRLD-----DSSQARDSTSY-----LTGEPKIEANGNHT 1643
QY 1432 SKLLPEKKSERSPLFOTDLKLKGGGLRYQKLPSDEDESGTGRVOITPHCSKMITRK 1491
DB 1644 AEVIPEAKAK-----PYEPESONDIGKQSIK-----ENLK 1673
QY 1492 AKORCASPOEHSABEIRTFIKAKELYSDALD--KKOSSDSGVRSNESSPNHSLHNEA 1549
DB 1674 PKTHGCGRTPE-PVSPLTAYOKSLEETSKLVIEDAKPCVPVGMK-----KMTRTTA 1724
QY 1550 DDSQLEKANLIELEDEGHSGKROMPHSLSGLDPTIARMSICSEDKS 1597
DB 1725 DG-----KARLNLQEEGST--RSEPQGGYK--VTKKEINRVEKKT 1764

```

Search completed: July 1, 2003, 14:42:36  
 Job time: 71 secs



•  
•  
•  
•

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:37:09 ; Search time 38 Seconds

(without alignments)  
4948.535 Million cell updates/sec

Title: US-10-021-571-2

Perfect score: 8853  
Sequence: 1 MSVLISQSIINVERENIPA.....ELHAASSTGCGEERSTL 1715

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published\_Applications\_AA.\*

1: /cgn2\_6/prodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/prodata/1/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/prodata/1/pubppa/US07\_PUBCOMB.pep:\*  
6: /cgn2\_6/prodata/1/pubppa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/1/pubppa/US08\_PUBCOMB.pep:\*  
8: /cgn2\_6/prodata/1/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/1/pubppa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/prodata/1/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/1/pubppa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/prodata/1/pubppa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/prodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2961.5	33.5	705	US-10-106-698-6378	Sequence 6378, App1
2	2782.5	31.4	551	US-09-835-788A-17	Sequence 17, App1
3	2260	25.5	513	US-10-149-819-9	Sequence 9, App1
4	549	6.2	109	US-09-986-480-371	Sequence 371, App1
5	538.5	6.1	1724	US-09-964-899-43	Sequence 43, App1
6	438	4.9	740	US-09-835-788A-12	Sequence 12, App1
7	401.5	4.5	426	US-09-908-711-70	Sequence 70, App1
8	395.5	4.5	1327	US-09-972-115A-8	Sequence 8, App1
9	395.5	4.5	1327	US-09-841-835-2	Sequence 2, App1
10	389.5	4.4	285	US-09-835-788A-18	Sequence 18, App1
11	387.5	4.4	949	US-09-841-835-10	Sequence 10, App1
12	380.5	4.3	1166	US-09-972-115A-6	Sequence 6, App1
13	380.5	4.3	1166	US-10-163-587A-15	Sequence 15, App1
14	371.5	4.2	1074	US-09-509-196A-2	Sequence 2, App1
15	360.5	4.1	1333	US-09-972-115A-2	Sequence 2, App1
16	360	4.1	784	US-10-164-080-7	Sequence 7, App1
17	355	4.0	251	US-09-835-788A-13	Sequence 13, App1
18	355	4.0	786	US-10-164-080-2	Sequence 2, App1
19	355	4.0	786	US-10-299-327-2	Sequence 2, App1

20	355	4.0	787	US-09-866-050A-334	Sequence 334, App1
21	354.5	4.0	1267	US-09-972-115A-4	Sequence 4, App1
22	354	4.0	802	US-09-964-899-41	Sequence 41, App1
23	349.5	3.9	679	US-10-339-936-2	Sequence 2, App1
24	331	3.7	833	US-09-947-199-2	Sequence 2, App1
25	331.5	3.6	673	US-09-841-835-8	Sequence 8, App1
26	303	3.4	835	US-09-947-199-8	Sequence 8, App1
27	299	3.4	306	US-09-835-788A-19	Sequence 19, App1
28	278	3.1	1719	US-10-012-896-378	Sequence 378, App1
29	278	3.1	1719	US-09-895-793-378	Sequence 378, App1
30	278	3.1	1719	US-09-895-814-378	Sequence 378, App1
31	278	3.1	1719	US-10-010-940-378	Sequence 378, App1
32	278	3.1	1719	US-09-759-143-378	Sequence 378, App1
33	278	3.1	1719	US-09-780-669-378	Sequence 378, App1
34	278	3.1	1719	US-09-822-827-378	Sequence 378, App1
35	263.5	3.0	599	US-09-735-368-2	Sequence 2, App1
36	249	2.8	622	US-10-197-666A-22	Sequence 22, App1
37	242	2.7	1054	US-09-798-042-87	Sequence 87, App1
38	240.5	2.7	622	US-10-197-666A-6	Sequence 6, App1
39	238.5	2.7	329	US-09-880-192-62	Sequence 62, App1
40	238.5	2.7	329	US-09-758-593A-1	Sequence 1, App1
41	236.5	2.7	328	US-09-758-593A-11	Sequence 11, App1
42	236.5	2.7	328	US-09-758-593A-12	Sequence 12, App1
43	236	2.7	319	US-10-312-054-1	Sequence 1, App1
44	233.5	2.6	210	US-10-127-032-167	Sequence 167, App1
45	233.5	2.6	460	US-09-905-673-66	Sequence 66, App1

## ALIGNMENTS

RESULT 1  
US-10-106-698-6378  
; Sequence 6378, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6378  
; LENGTH: 705  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (244)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (337)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; US-10-106-698-6378

Query Match 33.5%; Score 2961.5; DB 9; Length 705;  
Best Local Similarity 95.5%; Pred. No. 8.3e-181;  
Matches 567; Conservative 12; Mismatches 14; Indels 1; Gaps 1;

QY 62 ANCEDLDNWTALISASKEGHIHYEELKSGASLEHRDGGMTALMACYKGRDYVE 121  
DB 28 SMCNEDLDNWTALISASKEGHVHYEELKSGVNEHRDGGMTALMACYKGRDYVE 87  
QY 122 LLISHGANSYVGLGVSYPIIWAAGRGHADIYHLLONGAVNCSDKYGTPLVMAARK 181  
DB 88 LLISHGANSYVGLGVSYPIIWAAGRGHADIYHLLONGAVNCSDKYGTPLVMAARK 146

QY 182 GHECEVKKHLLAGADVDOEGANSMTALIVAVKGYTQSYKEITLKRNPVNLTDKQNTAL 241  
DB 147 GHECEVKKHLLAGADVDOEGANSMTALIVAVKGYTQSYKEITLKRNPVNLTDKQNTAL 206  
QY 242 MASKEGHEIYODLLDAGTYVNIIDRSQDVLIGAVGHEIYRALLQRYADIDINGQ 301  
DB 207 MASKEGHEIYODLLDAGTYVNIIDRSQDVLIGAVGHEIYRALLQRYADIDINGQ 266  
QY 302 DKRTLLYMAVEKGNATVRDILQCNPDTEICTKDETEPLKATKRNIEVEVLLDKAK 361  
DB 267 DKRTLLYMAVEKGNATVRDILQCNPDTEICTKDETEPLKATKRNIEVEVLLDKAK 326  
QY 362 VSAVDDKGDPLHVAIRGSRRLAELLRNPDGRLLYRPNKAGETPNIDCSHOKSLT 421  
DB 327 VSAVDDKGDPLHVAIRGSRRLAELLRNPDGRLLYRPNKAGETPNIDCSHOKSLT 386  
QY 422 QIFGARHLSPTETDDMDIGDLYSSALADILSEPTMQPPICVGLYQWGSKSFLLKLE 481  
DB 387 QIFGARHLSPTETDDMDIGDLYSSALADILSEPTMQPPICVGLYQWGSKSFLLKLE 446  
QY 482 DEMKTFAGQOETPELPFQFSMLYFLLCGGLGIVEAPVPDNLAIATISFLALITYEF 541  
DB 447 DEMKTFAGQOETPELPFQFSMLYFLLCGGLGIVEAPVPDNLAIATISFLALITYEF 506  
QY 542 IYIYFGRRGEGSMNMAWALSTRLARHIGYLELLFKLMFVNPPELPEOTTKALPVRFLE 601  
DB 507 IYIYFGRRGEGSMNMAWALSTRLARHIGYLELLFKLMFVNPPELPEOTTKALPVRFLE 566  
QY 602 DYNRLSSVGETSLAEMATISDACEREGFLATRLRFVFRTEBSQKKKKKK 655  
DB 567 DYNRLSSVGETSLAEMATISDACEREGFLATRLRFVFRTEBSQKKKKKK 620

## RESULT 2

US-09-835-788A-17  
; Sequence 17, Application US/09835788A  
; Patent No. US20020077458A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides,  
; FILE REFERENCE: P018P1  
; CURRENT APPLICATION NUMBER: US/09/835,788A  
; PRIOR FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: PCT/US00/28666  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/159,585  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 60/167,246  
; PRIOR FILING DATE: 1999-11-24  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 551  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-835-788A-17

Query Match 31.4%; Score 2782.5; DB 10; Length 551;  
Best Local Similarity 96.7%; Pred. No. 1.6e-169;  
Matches 534; Conservative 8; Mismatches 9; Indels 1; Gaps 1;  
QY 102 MGNWTALMAMCYKGRITDVVELLSHGANPSVTGLQYSVYPLIMAGRGHADIYHLLQNG 161  
DB 1 MGNWTALMAMCYKGRITDVVELLSHGANPSVTGL-YSVYPLIMAGRGHADIYHLLQNG 59  
QY 162 AKVNCSDKGYTTPLYMAARKGHECYVKKHLLAGADVDOEGANSMTALIVAVKGYTQSYK 221  
DB 60 AKVNCSDKGYTTPLYMAARKGHECYVKKHLLAGADVDOEGANSMTALIVAVKGYTQSYK 119  
QY 222 EILKRNPNVNLTDKQNTALMASKEGHEIYODLLDAGTYVNIIDRSQDVLIGAVRG 281

DB 120 EILKRNPNVNLTDKQNTALMASKEGHEIYODLLDAGTYVNIIDRSQDVLIGAVRG 179  
QY 282 HVEIVRALLQRYADIDIRGODNKTALYMAVEKGNATVRDILQCNPDTEICTKDETEPL 341  
DB 180 HVEIVRALLQRYADIDIRGODNKTALYMAVEKGNATVRDILQCNPDTEICTKDETEPL 239  
QY 342 KATKRNIEVEVLLDKGAVSAVDKGDPLHVAIRGSRRLAELLRNPDGRLLYRPN 401  
DB 240 KATKRNIEVEVLLDKGAVSAVDKGDPLHVAIRGSRRLAELLRNPDGRLLYRPN 299  
QY 402 NKAGETPNIDCSHOKSLTQIFGARHLSPTETDDMDIGDLYSSALADILSEPTMQPP 461  
DB 300 NKAGETPNIDCSHOKSLTQIFGARHLSPTETDDMDIGDLYSSALADILSEPTMQPP 359  
QY 462 CVGLYQWGSKSFLLKLEDEKKTAGQOETPELPFQFSMLYFLLCGGLGIVEAPVP 521  
DB 360 CVGLYQWGSKSFLLKLEDEKKTAGQOETPELPFQFSMLYFLLCGGLGIVEAPVP 419  
QY 522 DTNLAIATISFLALITYEFIVYFGRRGEGSMNMAWALSTRLARHIGYLELLFKLMFV 581  
DB 420 HPNLGIATISFLALITYEFIVYFGRRGEGSMNMAWALSTRLARHIGYLELLFKLMFV 479  
QY 582 NPPELPEOTTKALPVRFLETDYRLSSVGETSLAEMATISDACEREGFLATRLRFV 641  
DB 480 NPPELPEOTTKALPVRFLETDYRLSSVGETSLAEMATISDACEREGFLATRLRFV 539  
QY 642 RTEBSQKKKKKK 653  
DB 540 KTEDTQKKKKK 551

## RESULT 3

US-10-149-819-9  
; Sequence 9, Application US/10149819  
; Publication No. US20030044913A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: YUE, Henry  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LU, Dying Aina M.  
; APPLICANT: SHAH, Purvi  
; APPLICANT: TAL, Preethi  
; APPLICANT: AN-YOUNG, Janice  
; APPLICANT: BURFORD, Neil  
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES  
; FILE REFERENCE: PF-0760 PCT  
; CURRENT APPLICATION NUMBER: US/10/149,819  
; PRIOR FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354  
; PRIOR FILING DATE: 1999-12-10; 1999-12-16  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: misc\_feature  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20030044913A1 2215706CD1  
US-10-149-819-9

Query Match 25.5%; Score 2260; DB 9; Length 513;  
Best Local Similarity 86.6%; Pred. No. 3.3e-136;  
Matches 445; Conservative 21; Mismatches 46; Indels 2; Gaps 2;  
QY 1203 MAMNFGDMLFRSNTYLEMKSVEQVPEDEPRFLNENSSAPVPHGSRARSHTEPLTEL 1262  
DB 1 MAMNFGDMLFRSNTYLEMKSVEQVPEDEPRFLNENSSAPVPHGSRARSHTEPLTEL 60  
QY 1263 SSQPTTYLNFSEELNIGLDEGAPRHSNLSWQSQTRRTPTSLSSLNQSSIEISKLTDK 1322

```

      TITLE OF INVENTION: Identification of Genes Involved in
      TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
      FILE REFERENCE: 4-31612 A
      CURRENT APPLICATION NUMBER: US/09/964,899
      CURRENT FILING DATE: 2001-09-27
      PRIOR APPLICATION NUMBER: 60/236,893
      PRIOR FILING DATE: 2000-09-29
      PRIOR APPLICATION NUMBER: 60/298,309
      PRIOR FILING DATE: 2001-06-14
      NUMBER OF SEQ ID NOS: 53
      SOFTWARE: FASTSEQ for Windows Version 4.0
      SEQ ID NO 43
      LENGTH: 1724
      TYPE: PRT
      ORGANISM: Homo Sapien
US-09-964-899-43

Query Match      6 1%; Score 538.5; DB 9; Length 1724;
Best Local Similarity 20.3%; Pred. No. 2e-25;
Matches 390; Conservative 250; Mismatches 661; Indels 621; Gaps 72.

      14 EENIPAKALLEKCDVDERNECGQPTPLMAEAGNEIYVELLKNGANCNLDEDLDNMT 73
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      45 KEGHEVVESELLQRANPDATKKGNTALHLIASLAGQAEVVKLVLTNGANVNAQSQNFT 104
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      74 ALISASKEGHIHYELLKSGASLEHRDMGCGTALMAACVYGRDVEYELL----- 124
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      Db      105 PLYMAAQNHLEVVKFLIDNGASOSIATEDGFTPLAVALQOQHDDVVSLLLENDTKGRVR 164
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      QY      125 -----SHGANPSVTGL----- 135

```

[illegible]

```

QY 628 REFGE-----LATRLFRVTERESQKKKKMK-----TCC----- 658
  |||
Db 766 GYMGSFLARSARFLVSPVAVDARGSMGSRHGRITIPRCKTAPRTICRLKYRKKL 845
QY 659 -----PSVLEFLVGCIIAGITLLAIFRVDPKHLTVNAILISIA 698
  |||
Db 846 ANPPMVEGEGIASRLVEMGPAQOFL--CPVIEIPHFGSMGKREELIV-----LR 896
  |||
QY 699 SVVGLAFVLCNRTW-----OYLSLINSQKKRLHS-----AASKHLKSGF----- 742
  |||
Db 897 SE-----NGETKEHQFDSKNEDLTLLNGMDELDSPRELCKRIRCLITTDFOQYF 949
  |||
QY 743 --MKVLCEVELMARMAKTIDFTQNTRLVIID-----GLDACE--ODKV 785
  |||
Db 950 AVVSRKQESNOIGPEGILSST-----VPIVQSPREGALKRIRVGLQAPVDEI 1003
  |||
QY 786 LQMLDTVAVLSKGFIIASDHIITIKALINQNLVLRDSNNGHDYMRNIVHLPVL 845
  |||
Db 1004 VK-----KILGNKATFSPITVEPR-----RRKFKKPIYM 1033
  |||
QY 846 -----NSRGLSNARKFLVTSATNGDIT-----CSDTGTQ-----ED-----TDRVSQ 884
  |||
Db 1034 TIIPPPSGEVSNGYK-----GDTPLRLCLSTIGTSPAQMEDITGTPLTFIK 1085
  |||
QY 885 NSLGBMTKLSKTLNRRDYYRRQOMRTTIRMSFDLTKLVTEDWFSDISPQTMRL 944
  |||
Db 1086 DCVFTTVNSAR-----YGNKGQKAV-----LEGKPIVDCYGNLAPLTK----- 1126
  |||
QY 945 NIVSVTGLRLANOITFWMODLAWINL--TEQPIYTSMLIYLEETBGLPDMILKTM 1002
  |||
Db 1127 -----GGQOIVFNYSKRENRLPSIKIRDSQPCGRSLFKPKTKGILP-----QTA 1176
  |||
QY 1003 YERISKNIPTTKDVEPLEIDGDIRNEVELSRTPLVAVADYTFPCVINDPKREI 1062
  |||
Db 1177 VCNLNTILPAHKKIEKT-----DRR--QSFAS-----LALRKRSYLTGEMKEKKQSE 1224
  |||
QY 1063 IADVRAREQINIGLAVPLPLHEGPPRPESGTSQASVSSASFNGPPGGVSPQPH 1122
  |||
Db 1225 LSD-----EESTSRNYSLSSETSRG-----QPS 1248
  |||
QY 1123 SSYSYGLSGPHPPYNAVAVATGSSLLSSMTVDVYCEKIROEGDQMMMOYCTIK 1182
  |||
Db 1249 VTTKSARDKTEAPLKSSEKAGSEKRSRRTRADALTSVLTIKNRID-----IVTL- 1300
  |||
QY 1183 KANINGVYLSOCNIDELKKEMANFGDW-HLFRSMVLEMSVSQVVPEDPRELINENSSA 1241
  |||
Db 1301 --LEGPIFYGNISGTR-----SFADENNVFHDVPDGYPSL--QVELFPTGLHTPPT 1350
  |||
QY 1242 PVPHGESARSSHTLPLELSSQTPYTLNFSFEELNTLGLDGCARHSHLSMOSQTRRT 1301
  |||
Db 1351 PROODDYFSDISIESPL-----RTBSRLS-----DELVPQGNIEHSAD---G 1391
  |||
QY 1302 PSLSLSLNSODSIEISKLTDKVO--AEYRDA-----YREYIAQSOLEGGT 1345
  |||
Db 1392 PPV--VTAEDASLEDKLEDSVPLTEPRAVNDVDESOLENCLESEYFOYIGNIA----- 1443
  |||
QY 1346 GSSTISGRSSPHSTYIIGSSSGGSHSTLEDERGEG-----ELKQEDG 1390
  |||
Db 1444 GSPKDVKPAPEPR--LGVSSE-----QCEKGSQPEDEMEWEKLSLFDIEDLEG 1492
  |||
QY 1391 RKS-----FLMRGVIDYSSSGVS--TNEASP-----LDPITTE 1423
  |||
Db 1493 VESEBMTKEKVAILKRVQOALEMSSITGMONETSGNLESCAQRARYTGGLDLRLDS 1552
  |||
QY 1424 DEKSDSGSKLPLGKKSSEPSLFOYDLKLGGLRYOKLPSDEDSGSGVQYIPPHGSK 1483
  |||
Db 1553 PDGCRDSITSLYLGEGKPEANGSHETPEATKTY--FPESQNVG----- 1598
  |||
QY 1484 MITTKLKAQKRCASPOHS-----APITFTIKAEYLSDALI-----DKDS 1528
  |||
Db 1599 -----KOSTKETLKPKIHSQVHEPASPPLAAYOKSLETSKLIETKPCVPDLKNS 1651
  |||
QY 1529 -SDSG-----VRNNESS--PNHSLHNEADDSOLEKKNLILEDEGHS 1568
  |||

```

```

Db 1652 ESDDSSSEERRVTRVIRRLRLIKGEANNIPGESVTEBOFTD---EEGNLTRRGEGRK 1708
  |||
QY 1569 GK 1570
  |||
Db 1709 VK 1710
  |||

RESULT 6
US-09-835-788A-12
; Sequence 12, Application US/09835788A
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: Nt et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides
; FILE REFERENCE: PTO18P1
; CURRENT APPLICATION NUMBER: US/09/835,788A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-788A-12

Query Match      4.9%; Score 438; DB 10; Length 740;
Best Local Similarity 29.9%; Pred. No. 1.5e-19;
Matches 145; Conservative 67; Mismatches 193; Indels 80; Gaps 13;

QY 6 SOSVINYVEENIPPLKALLECKDYNDEBEGC--QTPMLAAEGGNVIVELLKNGNC 64
  |||
Db 303 SQTTSVNASQSMPPYPSV-----DIDATESNHMTALTLCAGGHEELVYLARDKTI 357
  |||
QY 65 NLEDLNTWALISASKEGHIHIVEELLKSGASLE--HRDWGWTALMAACYGRDVELL 123
  |||
Db 358 EHRDKKGFPLILATAGHVGVETLLDKGDIEAOSERTKDTPLSLACSGRGQEVDDL 417
  |||
QY 124 LSHGANPSVTGLQYSVYPIIMAGRGHDIYHILLQNAKYN--CSDKYGTPLPYMAARK 181
  |||
Db 418 LARGANKEHRNVS--DYTPLSLASAGGYNIITIKILLMAGAEINSRTGSKGISPLMLAAM 476
  |||
QY 182 GHLECVKHLAMGADV--QEGANSMTALIVAVGGYTSVKEILLKRNPNVNLTKDGNITA 240
  |||
Db 477 GHVPAVKLLDMGSDINNOIETNRNTALTLCFQGRAEVESLLIDKRANVHRKKTGITP 536
  |||
QY 241 LMIASKEGHIEIVODLDAGTYVNIIP--DRSGDYVLIGAVRGVETVIRALLQRYADIDI 298
  |||
Db 537 LMEASGGYAEYGVLLDKGADVNAIPVPSRRDALTLAADKGYKFCCELLIHRGAHIDV 596
  |||
QY 299 RGQDNKTLMAVAVENGATMTWRDILQCPDTEICTKQGEPLIKATKMRNIEVELLDK 358
  |||
Db 597 R-----NKGNNPMLMASNGHEDVQVLLVOA 623
  |||
QY 359 GAKVSAVDKGDTPPLVAVIRGRSRRLAELLRNPKDGRLLYRPKAGETPNIDCSHQKS 418
  |||
Db 624 GADVDAADNRKITPLMSFRKGHVKVVOYLK-----EVQFSDIEC----- 666
  |||
QY 419 ILTQIFGARHLSPETTDGMLGYDLYSSALADILSEPTMOPPLICVGLYAGWGSQKSPFLK 478
  |||
Db 667 -----MRVIA--TITDELL--KKHCQCVETIVKADQQA-----AEANKNASILTK 709
  |||
QY 479 KLEDE 483
  |||
Db 710 ELDE 714
  |||

```

[illegible]





QY 133 -----T 133  
Db 465 NCHGSAVDMAPFPELRELTFFKSHSLQAAREADLAKVKTALLETINKQPOSHET 524  
QY 134 GLOYSV-----YPIIWAAGRHADIVHLLONGARV 164  
Db 525 ALHCVAASLHPRKROYTELLRKGANVNEKNDFMPLHVAABRAHNDVMEYLHKRGAKM 584  
QY 165 NCSDKGTTPLVAAAKRHECYKHLAMGADVDGANSMTA----- 207  
Db 585 NALDTLGOTALHRAALAGHLQTCRLTLSTGSDPSITSLQGTAAQNGNEAVQOILSESP 644  
QY 208 -----LIVAKGGYTQSVKEILKRNPNVNLTKDG--NTALMAKKEGHEIYVDL 256  
Db 645 IRSDVDYRLLESKAGDLETVKQLS--SONVNCRLLEGHSHPLHRAAGYNVSVYVEL 703  
QY 257 LDAGTYVNPDRSGDVLGAVRGHVEIYRALLQKYADIDIRGQDNKTALVAVKGG-- 314  
Db 704 LHHGADVHAADKGLVPLHNAACSYGHVEVAELLVRRGASVYVADLWKFPLHBAAKGKY 763  
QY 315 -----NATWVRDILQ-----C 325  
Db 764 EICKLLIKHGADPTKKNRDNPTPLDYKESGDTIDILLKGDALLDAKGCCLARYQKLC 823  
QY 326 NPDTETCTKD---GEPPLIKATMRNIEVELLEDRGAKYSAVDKKDPLH----- 374  
Db 824 TPEINCRDTQGRNSPPLHLAAGYNNLEVAEYELLEHGADVNAQDKGLPLHNAASYGHV 883  
QY 375 -----VAIRGSRRLAELLRNPKGRLLYRNKAGETP 408  
Db 884 DIALALKYNTCVNADKMAFTPLHBAOKGRF-QICALLAHAGAPTM---KNQSGQTP 939  
QY 409 YN-----IDCSHOKSILT-----QIFGARHLSPTEPDGMLGYDLYSSALADI 451  
Db 940 LDATADDDIALLIDAMPRLPFTCKEPQATVVASLISAPSPS-----CLSAASSIDN 994  
QY 452 LSEPTWOPICVGLYAQWGS-----KSFLLKLEDEMKT 486  
Db 995 LTGRLAE--LAVGASNAAGGAGTERKEGEVAGLDNMNISQFLKSLGLEHLRIEFT 1049

RESULT 10  
US-09-835-788A-18  
; Sequence 18, Application US/09835788A  
; Patent No. US2002007458A1  
; GENERAL INFORMATION:  
; APPLICANT: NI et al.  
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, and  
; FILE REFERENCE: PTO18P1  
; CURRENT APPLICATION NUMBER: US/09/835, 788A  
; CURRENT FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: PCT/US00/28666  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/159,585  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 60/167,246  
; PRIOR FILING DATE: 1999-11-24  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 285  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-835-788A-18

Query Match 4.4%; Score 389.5; DB 10; Length 285;  
Best Local Similarity 37.8%; Pred. No. 4.2e-17;  
Matches 101; Conservative 49; Mismatches 110; Indels 7; Gaps 5;

QY 43 MAAEGNVEIVEKLLKNGANCNLEDD--NMTALISKSGCHIHIVELLKSGASLEHRD 101  
Db 1 MEASAGHVEVARVLLDHGAGINTHSNEFESALTTLACYGKHLDMVRFLEAGADOBHKT 60

QY 102 MCGWTALMACYKGRDVLVELLSHGANSVTGLQYSVYPIIWAAGRHADIVHLLONG 161  
Db 61 DEMHTALMBACMDGHVEVARLLDLSGAGVNNPDSF--SPITLAAAGSHVLAALLIERG 119  
QY 162 AKVNCSDRYGTPPLVMAARKHLECYKHLAMGADV--QEGANSMTALIVAKGYTQSV 220  
Db 120 ANLEEVNDEGYPLMEAREGHEEVALLILAQGANINNOTBEQETALTLACGGSFEVA 179  
QY 221 KEILKRNPNVNLTKDGNTALMAKKEGHEIYVDLDTAGTYVNPDRSGDVLIGAVRG 280  
Db 180 DELIKAGADIEL---GCTPLMEASQEGHLELVKYLKSLASGANVATTAATGDTALTACEN 236  
QY 281 GHEIYRALLQKYADIDIRGQDNKTAL 307  
Db 237 GHTVDADVLLQAGADD--KQEDMKITL 262

RESULT 11  
US-09-841-835-10  
; Sequence 10, Application US/09841835  
; Patent No. US20020076795A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Tilia  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,387  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 949 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-841-835-10

Query Match 4.4%; Score 387.5; DB 10; Length 949;  
Best Local Similarity 25.2%; Pred. No. 3.6e-16;  
Matches 144; Conservative 59; Mismatches 177; Indels 191; Gaps 11;

QY 14 EENETPALKALECKDVERNECGOTPLMLAEGNVEIVEKLLKNGANCNLEDDNMT 73  
Db 348 EEKIMALTPPLANNCHASDGRK---STPLHLAAGYVRVIVOLLQHGADVHAADKGLV 404  
QY 74 ALISAKKEGHIHIVEELKSGASLEHRDMCGWTALMACYKGRDVLVELLSHGANSVS- 132

Db 405 PLHNACSYGHEVTELLKHGACVANMDLMOFTPLHEAASKNREVECSLLSHGADPTLV 464  
QY 133 -----  
Db 465 NCHGSAVDMAPTEPLERLELYEFKGSHLQAAAREADLAKYKKTALLETINFKOPOSHET 524  
QY 134 GLOYSV-----YPIIWAAGRGHADVHLLONGAKV 164  
Db 525 ALHCAVASLHPKRGQVTELLKHGANVNEKNDEFTPLHVAERAHNDVMEYLKHGAKM 584  
QY 165 NCSDKYGTPLVMAARKGHECVKHLAMGADVDOEGANSMTA-----207  
Db 585 NALDPLGOTLHRAALAGHLQTCRLLSYGSDPSIISLOGFTAAQMGNEAVOQILSESTP 644  
QY 208 -----LIYAAGGYTOSYKELKRNPNVNLTDKCG--NTALMISKEGHEIYVDL 256  
Db 645 IRTSDVDRLLERASAGDELEVQJCS--SQNVNCDLEGRHSTPLHFAAGYNRVSVEYL 703  
QY 257 IDAGTYVNIPIRSGDTVLIGAVRGHEIVRALLOKADIDIRGODNKTALYMAVEKG--314  
Db 704 LHHGADVHAARKGGLVPLHNACSYGHEVALLVHGASVAVADMKTFTPLHEAAAKGY 763  
QY 315 -----NATVARDLIQ-----C 325  
Db 764 EICKLLKHGADPTKKNBDGNTPLDVKEGDIDLDLKGDAALLDAAKKGLARVOKLC 823  
QY 326 NPDTEICTKD--GETPIKATKMRNIEVEVLLDKGAKYSAVDKGGDTPLHVAIRGRSR 382  
Db 824 TPENINCDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVANAQDKGLIPLHNA-----877  
QY 383 RLAEILLRNPDKDRLLYRPNKAGETPNYIDC 413  
Db 878 -----ASYGGLARVOKLC--TPENINC 898

RESULT 12  
US-09-972-115A-6  
; Sequence 6, Application US/09972115A  
; Publication No. US20030032769A1  
; GENERAL INFORMATION:  
; APPLICANT: Geron Corporation  
; APPLICANT: Gregg, Morin B.  
; APPLICANT: Walter, Funk D.  
; APPLICANT: Mieczyslaw, Platysek A.  
; TITLE OF INVENTION: A Second Mammalian Telomerase  
; FILE REFERENCE: 080/003C  
; CURRENT APPLICATION NUMBER: US/09/972,115A  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: US 60/128,577  
; PRIOR FILING DATE: 2000-04-10  
; PRIOR APPLICATION NUMBER: US 60/129,123  
; PRIOR FILING DATE: 1999-04-13  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 1166  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-972-115A-6

Query Match 4.3%; Score 380.5; DB 9; Length 1166;  
Best Local Similarity 24.0%; Pred. No. 1,4e-15;  
Matches 157; Conservative 73; Mismatches 206; Indels 217; Gaps 13;  
QY 14 EENIPALKALLECKDVERNECGOTPLMLAEOGNEVIELKNGANCNLEDLDMWT 73  
Db 190 EEMKALLTPLNVNCHASDGRK---STPLHLAAGYNRKYVOLLLOHGADVHAARKGDLV 246  
QY 74 ALISASKEGHIHVEELKSGASLEHRDQGTALMAMCYKGRDVTVELLSHGANSPTV 133  
Db 247 PLHNACSYGHEVTELLKHGACVANMDLMOFTPLHEAASKNREVECSLLSYGADPTLV 306  
QY 134 G-----LQYS-----138

Db 307 NCHKSAIDLAPTPOLKERLALYEFKGSHLQAAAREADYTRIKKHSLEWNEKHPOTHET 366  
QY 139 -----YPIIWAAGRGHADVHLLONGAKV 164  
Db 367 ALHCAASPYKRGQJCELLKHGANINEKTEFELPLHVASKANDVVEYVVKHAKV 426  
QY 165 NCSDKYGTPLVMAARKGHECVKHLAMGADV-----OEG--201  
Db 427 NALDNLGOTSLHRAACGHLQTCRLLSYGSDPSIISLOGFTALQMGNEVVOQLQEGIS 486  
QY 202 -----ANSMTALIVAAGGYTOSYK-----TL 224  
Db 487 LGNSEADROLLEAAKAGDVEYVVKCTVGSVNCRDIEGROSTPLHFAAGYNRVSVEYL 546  
QY 225 KRNPNVNLTKDGNLTALMISKEGHEIYVDLIDAGTYVNIPIRSGDTVLIGAVRGHE 284  
Db 547 QHGADVHAARKGGLVPLHNACSYGHEVALLVHGAVVAVADMKTFTPLHEAAAKGYE 606  
QY 285 IVRALLQKADIDIRGODNKTALYMAVEKGNATVARDLIQ-----324  
Db 607 ICKLLLOHGADPTKKNBDGNTPL--DLVKGDIDIDLLRGDAALLDAAKKGLARVOKLC 664  
QY 325 CNPDTEICTKDE--TPIKATKMRNIEVEVLLDKGAKYSAVDKGGDTPLH-----374  
Db 665 SSPDNVNC--RDQGRHSTPLHLAAGYNNLEVAEYLLOHGADVANAQDKGLIPLHNAASYG 723  
QY 375 -----VAIRGRSRLAEILLRNPDKDRLLYRPNKAGE 406  
Db 724 HVDVAALLIKYNACVANATDKMAFTPLHEAAOKRT--QLCALHLAHGADPTL--KNQEOG 779  
QY 407 TPYNIDCSHOKSITQIFGARHLSPTETODMGYULYSALADIISEPTMP 459  
Db 780 TPIDIVSADVSAALITRAAMPRLSPSCYKPVQVINGVRSFGATRALISSPSSP 832

RESULT 13  
US-10-163-587A-15  
; Sequence 15, Application US/10163587A  
; Publication No. US20030096263A1  
; GENERAL INFORMATION:  
; APPLICANT: Oliveira, Marcos  
; TITLE OF INVENTION: SELECTIVE PARP-1 TARGETING FOR DESIGNING CHEMO/RADIO SENSITIZI  
; FILE REFERENCE: 50229-306  
; CURRENT APPLICATION NUMBER: US/10/163,587A  
; PRIOR FILING DATE: 2003-01-10  
; PRIOR APPLICATION NUMBER: 60/236,110  
; PRIOR FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 1166  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-163-587A-15

Query Match 4.3%; Score 380.5; DB 9; Length 1166;  
Best Local Similarity 24.0%; Pred. No. 1,4e-15;  
Matches 157; Conservative 73; Mismatches 206; Indels 217; Gaps 13;  
QY 14 EENIPALKALLECKDVERNECGOTPLMLAEOGNEVIELKNGANCNLEDLDMWT 73  
Db 190 EEMKALLTPLNVNCHASDGRK---STPLHLAAGYNRKYVOLLLOHGADVHAARKGDLV 246  
QY 74 ALISASKEGHIHVEELKSGASLEHRDQGTALMAMCYKGRDVTVELLSHGANSPTV 133  
Db 247 PLHNACSYGHEVTELLKHGACVANMDLMOFTPLHEAASKNREVECSLLSYGADPTLV 306  
QY 134 G-----LQYS-----138  
QY 307 NCHKSAIDLAPTPOLKERLALYEFKGSHLQAAAREADYTRIKKHSLEWNEKHPOTHET 366  
QY 139 -----YPIIWAAGRGHADVHLLONGAKV 164

```

Db 367 ALHCAASPYPRKQICELLRKGANINKEKTEFLPLHVASSEKANDVEYVYKBEAKV 426
QY 165 NCSDKYGTPLVMAARKGHLECYKHLAMGADV-----QEG-- 201
Db 427 MALDNLGOTSLHRAAYCGHLOTGCRLLSYGCDPNIISLGFTALQMGNNVQQLQEGIS 486
QY 202 ---ANSMTALIVAVKGYTSYKVE-----IL 224
Db 487 LGNSEADROLLEAKAGDVEYKLCITYQSVNCRDIEGRSTPLHFAAGYNRSYVEYLL 546
QY 225 KRPNVNLTDKDGNTALMTASKSGHIEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVE 284
Db 547 QHGADVHAKDKGIVPLHNAACSYGHEVAELLYKGAIVNVADLMKFTPLHAAAKGYE 606
QY 285 IYRALLQKADIDIRGQDNKTALYMAVEKGNATVPRDILQ----- 324
Db 607 ICKLLQHGADPTKKNRDNTPPL--DLVKDGDPTDIDLLRGDAALLDAAKKGLARVKKL 664
QY 325 CNPDTEICTKDE---TPLIKATKMRNIEVEVELLDGAKVASAVDKKGDPTPLH----- 374
Db 665 SSPDNVNC-RDYGGRHSTPLHLAGYNNLEVAEYLLQHGADVNAADKGLIPLHNAASYG 723
QY 375 -----VAIRGRSRRLAELLRNPKDGRLLYRPKNAGE 406
Db 724 HVDVALLIKYNACVNATDKMAFTPLHAAQKGR-QLCALLLHAGADPTL---KNQEGQ 779
QY 407 TPNYIDCSHOKSILTOIFGARHLSPTETDGMIGYDLYSSALADLSEPTMOP 459
Db 780 TPDLVVSADVVSALLTPAAMPSPALPSCYKQVNLNGVRSPGATDALSSGSPSP 832

```

```

RESULT 14
US-09-509-196A-2
; Sequence 2, Application US/09509196A
; Patent No. US20020037582A1
; GENERAL INFORMATION:
; APPLICANT: DALY, Roger J.
; APPLICANT: SOTHELAND, Robert L.
; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling
; FILE REFERENCE: 1871-129
; CURRENT APPLICATION NUMBER: US/09/509,196A
; CURRENT FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: P09388
; PRIOR FILING DATE: 1997-09-23
; PRIOR APPLICATION NUMBER: PCT AU98/00795
; PRIOR FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1074
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-196A-2

```

```

Query Match 4.2%; Score 371.5; DB 10; Length 1074;
Best Local Similarity 23.7%; Pred. No. 4.5e-15;
Matches 155; Conservative 73; Mismatches 208; Indels 217; Gaps 13;
QY 14 EENINIRALKALEKCDVBERNCCGOTPLMLAEGGNVEYKELLKNGANCNLEDIDNNT 73
Db 98 EEMMALLPPLVNCNASHDGRK---STPLHLAGYNNVIVDILQHGADVNAKKDGLV 154
QY 74 ALISASKEGHIHVEELKSGASLEHRDGMGTALMMACYGKGTVDVVELLSHGANSYST 133
Db 155 PLHNACSYGHEVTELLVHGGCVNADMLOFTPLHAAKSNVEVCSLLSYGADPTLL 214
QY 134 G-----LQYS----- 138
Db 215 NCKNSAIDLAFTPOLKERLAYEFKSHLSLQARADVTYRIKKHLSLEWVNFKNPQTHET 274
QY 139 -----VPIIWAAGRGADIVHLLLQNGAKV 164

```

```

Db 275 ALHCAASPYPRKQICELLRKGANINKEKTEFLPLHVASSEKANDVEYVYKBEAKV 334
QY 165 NCSDKYGTPLVMAARKGHLECYKHLAMGADV-----QEG-- 201
Db 335 MALDNLGOTSLHRAAYCGHLOTGCRLLSYGCDPNIISLGFTALQMGNNVQQLQEGIS 394
QY 202 ---ANSMTALIVAVKGYTSYKVE-----IL 224
Db 395 LGNSEADROLLEAKAGDVEYKLCITYQSVNCRDIEGRSTPLHFAAGYNRSYVEYLL 454
QY 225 KRPNVNLTDKDGNTALMTASKSGHIEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVE 284
Db 455 QHGADVHAKDKGIVPLHNAACSYGHEVAELLYKGAIVNVADLMKFTPLHAAAKGYE 514
QY 285 IYRALLQKADIDIRGQDNKTALYMAVEKGNATVPRDILQ----- 324
Db 515 ICKLLQHGADPTKKNRDNTPPL--DLVKDGDPTDIDLLRGDAALLDAAKKGLARVKKL 572
QY 325 CNPDTEICTKDE---TPLIKATKMRNIEVEVELLDGAKVASAVDKKGDPTPLH----- 374
Db 573 SSPDNVNC-RDYGGRHSTPLHLAGYNNLEVAEYLLQHGADVNAADKGLIPLHNAASYG 631
QY 375 -----VAIRGRSRRLAELLRNPKDGRLLYRPKNAGE 406
Db 632 HVDVALLIKYNASINATDKMAFTPLHAAQKGR-QLCALLLHAGADPTL---KNQEGQ 687
QY 407 TPNYIDCSHOKSILTOIFGARHLSPTETDGMIGYDLYSSALADLSEPTMOP 459
Db 688 TPDLVVSADVVSALLTPAAMPSPALPSCYKQVNLNGVRSPGATDALSSGSPSP 740

```

```

RESULT 15
US-09-972-115A-2
; Sequence 2, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Genon Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Platyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: The 'xaa' at location 1 stands for Lys, Asn, Arg, Ser, Thr, Il
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon
; OTHER INFORMATION: yr, Trp, Cys, or Phe.
; NAME/KEY: misc.feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: The 'xaa' at location 2 stands for Lys, Asn, Arg, Ser, Thr, Il
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon
; OTHER INFORMATION: yr, Trp, Cys, or Phe.
; NAME/KEY: (3)..(3)
; OTHER INFORMATION: The 'xaa' at location 3 stands for Lys, Asn, Arg, Ser, Thr, Il
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon
; OTHER INFORMATION: yr, Trp, Cys, or Phe.
; NAME/KEY: (4)..(4)
; LOCATION: (4)..(4)

```









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:05 ; Search time 19.5 Seconds  
(without alignments)  
2587.708 Million cell updates/sec

Title: US-10-021-571-2

Sequence: 8853

1 MSVLISQSVINVEENIPAA.....ELHAASSTGFGFERESTL 1715

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Database: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
7: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
8: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
9: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
10: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
11: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
12: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
13: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
14: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
15: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
16: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
17: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
18: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
19: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
20: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
21: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
22: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
23: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
24: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
25: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
26: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
27: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	536	6.1	1839	2	US-09-172-977-4
2	503.5	5.7	843	2	US-09-172-977-3
3	466.5	5.3	1745	2	US-09-031-485-33
4	466.5	5.3	1745	2	US-08-847-429A-33
5	466.5	5.3	1745	3	US-09-065-474-33
6	466.5	5.3	1745	3	US-09-065-474-33
7	449.5	5.1	1088	4	US-09-082-059-2
8	414.5	4.7	1423	4	US-08-810-712-10
9	395.5	4.5	1327	4	US-09-196-387-2
10	387.5	4.4	949	4	US-09-196-387-10
11	376	4.2	348	2	US-09-031-485-28
12	376	4.2	348	2	US-08-847-429A-28
13	376	4.2	348	3	US-09-065-474-28
14	376	4.2	348	4	US-09-557-034-28
15	373	4.2	1166	4	US-09-350-982C-5
16	359.5	4.1	352	3	US-09-065-474-139
17	359.5	4.1	352	3	US-09-557-034-139
18	357.5	4.0	741	2	US-08-436-771-4
19	357.5	4.0	741	2	US-08-436-998-4
20	357.5	4.0	741	2	US-08-487-797-4
21	357.5	4.0	741	2	US-08-487-797-4
22	356	4.0	741	5	PCT-US95-02058-4
23	356	4.0	679	2	US-08-436-771-6
24	356	4.0	679	2	US-08-434-998-6
25	356	4.0	679	2	US-08-487-797-6
26	356	4.0	679	2	US-08-701-005A-4
27	356	4.0	679	2	US-08-479-895-4

28	356	4.0	679	5	PCT-US95-02058-6	Sequence 6, App1
29	355	4.0	787	4	US-09-188-930-334	Sequence 38, App
30	353	4.0	302	2	US-09-031-485-38	Sequence 33, App1
31	353	4.0	302	2	US-08-847-429A-38	Sequence 38, App1
32	353	4.0	302	3	US-09-065-474-38	Sequence 38, App1
33	353	4.0	302	3	US-09-557-034-38	Sequence 38, App1
34	353	4.0	303	2	US-09-031-485-23	Sequence 23, App1
35	353	4.0	303	2	US-08-847-429A-23	Sequence 23, App1
36	353	4.0	303	3	US-09-065-474-23	Sequence 23, App1
37	353	4.0	303	3	US-09-557-034-23	Sequence 23, App1
38	352	4.0	679	3	US-08-943-956A-4	Sequence 4, App1
39	347.5	3.9	741	2	US-08-462-481-2	Sequence 2, App1
40	347.5	3.9	741	2	US-08-436-771-2	Sequence 2, App1
41	347.5	3.9	741	2	US-08-434-998-2	Sequence 2, App1
42	347.5	3.9	741	2	US-08-487-797-2	Sequence 2, App1
43	347.5	3.9	741	2	US-08-701-005A-2	Sequence 2, App1
44	347.5	3.9	741	5	PCT-US95-02058-2	Sequence 2, App1
45	347.5	3.9	741	5	PCT-US95-02058-2	Sequence 2, App1

#### ALIGNMENTS

RESULT 1  
US-09-172-977-4  
; Sequence 4, Application US/09172977  
; Patent No. 5989663  
; GENERAL INFORMATION:  
; APPLICANT: Tangu, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Yue, Henry  
; TITLE OF INVENTION: HUMAN ANKRYN FAMILY PROTEIN  
; FILE REFERENCE: PF-0615 US  
; CURRENT APPLICATION NUMBER: US/09/172,977  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 1839  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: g29491  
US-09-172-977-4

Query Match	6.1%	Score 536;	DB 2;	Length 1839;
Best Local Similarity	19.7%	Pred. No. 1.4e-35;		
Matches 398;	Conservative 276;	Mismatches 706;	Indels 636;	Gaps 70;
14	EEENIPALKALKECKDYDERNEGQTPMLAEOGNEIYKELLKNGANCLEDDIMNT	73		
73	KSGHGLVQELLGRSSVDSATKRCNTALHLSLAGQAEVYKLVKEGANTNAGSONGT	132		
74	ALISAKGKHIVIBELLKSGAS	103		
133	PLYMAOENHIDVYKLLNGANOSTATEDGFTPLAVALOOGHNOVAALLLENPTKQVR	192		
104		123		
193	LPALHIAKDDTKSAALLLONHDNADYOSKMMVNRTEGTPHIAHAGNVAVTL	252		
124	LSHGANPSVTGQYVYPLIMAGRGHADIYHLLONCAKAYKSDKYTTPLVMAARKG	182		
253	LNRGAADVTA-RNGITPLHVASKRGNTNMVYLLDRGGQIDAKTRDGLTLHCAARSG	311		
183		210		
312	DQVVELLERGAPLARTKNGISPLHMAAQGDHVCVHLLHOKPARDVDVTLDTLVH	371		
211	AKVGQGVQVKKILKRNPNVNLTDKDGNTALMAISKE	247		
372	AAHGHYRVTKLLDKNRANPARNALNGFTPLHIAKKNRIKVMELLYKASIQAITESG	431		

```

248 -----GHIEIYODLLDAGTYVNIIPDRSGDTVLIGAVGGHVEIIRALLKXYADID 297
Db 432 LPIIHVAAPMGILNIVLILLONGASPDVNTIRGETALMAARAGAVEYVRCGLRCAVD 491
QY 298 IRGDNKTAALYWAVERGNATVNDILQCNPTLEICTKDEGTEPLKATMARIIEVELLD 357
Db 492 AARAREQTPLHIASRLGKTEIYQILLQMAHPDADATNGTYPHLSAEGGVDAVSLE 551
QY 358 KGAKSAVDDKGDPTLHVAIRGRSLRLABELLR-----NPKD 394
Db 552 AGAHSLSATKKGEPPLHVAARYGSLDVAKLILLORRAADASAGKNGLPPLHVAHYDMQV 611
QY 395 GRIL-----YRPKAGETPYNIDC--SHOKSILTOI--FGAR-----HLSPT 432
Db 612 ALLLEKASPRATKNGTYTPLHIAKKNQMOIASITLNTYGAENIYTKOGVTPHLASQ 671
QY 433 EIDGML-----GYDLYSALADILSEPTMOPICVGLYQWGSGLLKLEDEM-- 484
Db 672 EGHDMVTLILLDKGANIMHSTRSGILTS-----LHLAAQ-----EDKYNV 710
QY 485 -----KTFAGOOTEPLEFQSMILYFLLLLC--GGILYV--PAPVDTMLAIALSLPLAL 536
Db 711 ADILFKGADDAHFKLGITPLIV-----ACHYGVKKNVNFLLKQGANVNAKTNGYTPL 765
QY 537 -----IYFFIYIYFGRREGESNMAMALSTRLARHIGYLELLFKLMFVNP--- 584
Db 766 HOAAQOQHTHIIIVLLOHGAKPNATTTANGNTALA--IAKRGYISVDTLKVYIEVYTT 823
QY 585 -----ELPEOTYALPYRFLTYDYNLSVGGETSIAEMITAILSDACEREGPL- 633
Db 824 TTTTEKHKNLVPETMEVLVDVSEEDG-DTMTGDGEYLRPEDLKELGDSLPSQFLD 882
QY 634 -----ATRLFRVPRF-----EE 645
Db 883 GMYIARYLEGGRSLSRFSFSDSHILSHASYLSDANVDDSVYIRSHOVSTLAKAER 942
QY 646 SOKKKK-----KTYCCLPSFVIFLTY-----GCIITGLTLAIFRVDPKHL 688
Db 943 NSYRLSWGTEMLDNVALSSPISHGFLYIFWVDARGAMRCGRHNGRLRI--IPPRKC 998
QY 689 TVNALILISAVGIAFLNCRTMQVYLSLNSQKRL-----HSAKSLHLKLS 739
Db 999 TAPF-----RYTCRL-----VKRHRRLMPMVEGEGELRLIEVGP 1035
QY 740 EG-----FMVLKCEVELAMARKATIDSEFON-----OTRLVYIIDGLACE 781
Db 1036 SGAGFLGPVIEIHFALARKERELVYLRSENDSKHEKCDYTELEINELNGMD--- 1092
QY 782 QDKVLQMLDIYRVLEPSGPTAIPASDPH--IIRKAINQNLNSVLRSDNINHDMYRNV 839
Db 1093 -----EVLDSPEDELEKRRICRIITRDPQYFAVYSRIQ-----DSNLIGPE-- 1134
QY 840 HLPFLNLSRGLSNARKFLVYSAT--NGDITCSDTTQOEDTDRVSONSLGEMTK--LGS 895
Db 1135 -----GVLSTSTVYPOVOAYPEBALTKRLIVGQ-----AOPMHESELVKLLGN 1179
QY 896 KTALENRDYV--RRRQORITTRQMSFDLTKLVTEDMFSDISPOTKRRLNIYSVTRL 953
Db 1180 KATESPIVLEPRRRKHFKPTMTIIPVKKASDVMLNGFGDAP-TIARLL--CSITGCT 1235
QY 954 LRANQITFNDRLASWINLT--EOMPYRTS--WLIYLEETEGLPDMQTLKTYER 1005
Db 1236 TPA-----QWEDITGTPLEFVNCEVSTTNVSAFWLI-----DCRQIQESVTFASQVYR 1286
QY 1006 ISKNIPTTKDVEPLELTDGDIRNEFVLSRTPLVARADVTKPLPCVYNIDPKL----- 1059
Db 1287 EITCVPY-----MAKTVFAKSHD-----TEARIRCCMTDDVDVKTLEQ 1327
QY 1060 REIADYRAAREQINIGLAVPPLPLHECPPPRPPSGYSPASVCSASFNGPPGAVSP 1119
Db 1328 QENAEVARSND-----VEVLEGRPIYVDCGNLVPLTKSG----- 1363

```

```

QY 1120 QPHSSYSGLSGPHFPYNNRAVPATGSSLLSSM--TVDVCEKL-----ROIISLDON 1172
Db 1364 -----QHHSFPAFKENRLPLEFVAVRDTTOEPCGRSLSEKPEKSTGLVHQ 1410
QY 1173 MPOYCTTIKANINRVLISOCNIDELKREMANFDMHLFRSMVLEKMSVSQVPEDP 1232
Db 1411 -----ALONLNTIPLTKESNOQE-----EELDMTSKKNPOE 1447
QY 1233 RFLNENSAVPVH-----GESARRSSHLEPLTELSSOTPYTLNFSPELNTGLDEGA 1286
Db 1448 QERIEERLAIADHLGFSWTEIARELDFTEEQIHQIRINPNLSLODOSYLKIKWLERBG 1507
QY 1287 PRHSNLSWOSQTRPTSLNSINGDSISFESKLTLDVOAEYDAYEYIAQMSOLEGGTG 1346
Db 1508 -KIA-----TDVLVECLTKRINMDLVHLMETNTEPLOERISHSYAE--TEQITLIDHSEG 1560
QY 1347 SSTISGR-----SSPHSYIIGQSSSGSISHT 1374
Db 1561 FSVLOEELCTAQHKQKEQAVSKESBTDHPPIVSEBISVGSFTQDGVPRTEGDSST 1620
QY 1375 LEQERKEGELKQEDGRKSPIMKRGVIDIYSSGVSNTASPLDPTTEDEKSDSGSKL 1434
Db 1621 ALFPQTHKEQVODFSGKMODLPRESSLEYQOEYVT-----TPGETSETOKAMI 1671
QY 1435 LPG--KKSSEPSLFTQDLKLGGLRYOKLPSDESGTGRVQITPHCSKMIPTKRLKA 1492
Db 1672 VPSSPKTEPREVSTPAEKKL-----YLOTPSSRSGSPIIIO----- 1709
QY 1493 KORECASPOHSAE--PIRTFIKAEYLSDALDKKSDSGVRSNNESSPNHSLNEAD 1550
Db 1710 ---EPEPSHRESSPRKTSIYIV-----SADNOPENCERLDEDAAEKGD 1755
QY 1551 DSQLEKANLIE--LEDEHSGSGRGMPSHLSGLDPTIARMSICEDKKSPECCLIASS 1607
Db 1756 MPETPEYTEEYIDEHGHYVKKVTKR-----IIRRVSSBG--TEREELTVQOM 1805
QY 1608 PEESSMPA--COKAYNLNRPTSTYTLNNNTAPTNRANO 1642
Db 1806 PQE--PVNIEEGDYSKVIKRVYLSKDTQSDNNE 1839

RESULT 2
US-09-172-977-3
; Sequence 3, Application us/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. TOM
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: g1841966
US-09-172-977-3

Query Match 5.7%; Score 503.5; DB 2; Length 843;
Best Local Similarity 24.6%; Pred. No. 1,8e-33;
Matches 203; Conservative 110; Mismatches 273; Indels 239; Gaps 20;

QY 14 EENIPALKALKKQDVERNECGOTPLMLAEOGNEIYKELLKNGANGLDLDQMT 73
Db 38 KEGHGVIGVLLGRSSVDSATKKNATLHLSLAGQAEVYKLVKEGANIMASQNGFT 97
QY 74 ALISASKEGHIIHVEELKSGAS-----LEHRDMG--- 103

```

```

Db 98 PLYMAQENHIDVYKLLNGANQSTATEDGFTPLAVALQGHNOAVALLENDKGRK 157
QY 104 -----GFTALMACYKRTDVELL 123
Db 158 LPLAHIAARKDDTKSALLQDNHNAVQSKMMVNTTESGFTPLHIAHNGVNAVTL 217
QY 124 LSHGANPSVYGLQSVYPTIIMAGRGHADIYHLLQNGKAVCSQYGTPTIYMAARKG 182
Db 218 LNRGAADVPTA-RNGITPLHVASRKNMTVMVKLLDRGQIDAKTRDGLTPLHCAARGH 276
QY 183 -----HLECVKHLAMGADVDEGANSTALIV 210
Db 277 DQVVELLERGAPLARTKNGSLPLHMAQGHVEYKHLQHKAPVDVDTLDYLTALIV 336
QY 211 AVKGYTQSVKELKRNPNVNTDKDGNALMASKE----- 247
Db 337 AAHGHRVYTKLLDRANPNARALNGFTPLHACKKRIKVELLYKGYATQAITESG 396
QY 248 -----GHIEIYODLLDAGTYVNIIDRSQDYLIGAVGGHVEIYRALLQKADID 297
Db 397 LPTIYPAATMGHINITYLLQNASPDVTINIKETALHMAARGVEYVRCULRNGALVD 456
QY 298 IRGQDKTALYMAVEKGNATMVRDILQCNPDTEICTKDETPLIKATKRNIEVELLD 357
Db 457 ARAREGOTPLHIASRIGKTEIYQLLQHMAHPDAATNGVTPHISAREGYDVASVLE 516
QY 358 KGAKYAVNDKGDPTPLHVAIRGRSRLAELLIR-----NPRD 394
Db 517 AGAHSLATRKGTPLHVAAKYGSILYAKLLQRAADASAGKNGITPLHVAHAYDNOFY 576
QY 395 GRLL-----YRPNKAGETPNYND-C-SHOKSILTOI-FGAR-----HLSPT 432
Db 577 ALLLEKGSAPATAKNGITPLHIAKKKQMOJASTILNYGAFNTVNTQGVTPHLSAQ 636
QY 433 ETDGMLGYDLYSSALADILSEPTMOPICVGLYAWGSGKSF-LKLEDEM----- 484
Db 637 EGHDMVTLVLEKGA-----NIMHSTKSGLTSLHIAEDKYNVADILT 680
QY 485 KTAGAGOTPLPFPMSLIVPLTLLC--GGGLGV-FAFPVDTMLAIAISLAL----- 536
Db 681 KHGADDDATYTKLGYTPLIV---ACHYGNVKKVNFLLKOGAVNNAKNGYTPHQAQ 735
QY 537 ---IYFEIYIYFGREGESMMAMALSTRLARHIGYLELLFKLAFVNP----- 584
Db 736 QGRHIIINVLQGAKRNATTANGNTALA--IAKRLGYSVDTPLAVVEYTTTTTIT 793
QY 585 -----ELPEQTKALPYRFLFTDYNRLSSVGETSLAEMITATLSD 624
Db 794 EKHLNAPETMTETVLDSDEGD-DVYTGDSGEYLRPEDLKELGD 837

```

```

RESULT 3
US-09-031-485-33
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blum, E. Scott
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-485-33
;
; Query Match
; Best Local Similarity 20.4%; Score 466.5; DB 2; Length 1745;
; Matches 358; Conservative 235; Mismatches 575; Indels 591; Gaps 73;
;
; 14 EENIPALKALLEKCKDYDERNEGOTPLMAEAGNVEIYKELKNGANCLELDN-- 71
; 207 KDDPTKATLTLQNHNSDVTSGFTPLHIAHNGVNAVQALLLEKGANVYARHNIS 266
; 72 -----W-----TALISKRGHIIYELKSGASL-- 97
; 267 PLHATKGRIMVSLLAHGAVIDCRTRDILTLPLHCKSRGHDQVDLLEKAPISAK 326
; 98 -----EHRMGWTLAMW 110
; 327 TNGIAPLHMAQVDVYDITPLHVAHGHVAVAKLLDRNADPARALNGTPLHI 386
; 111 ACYKRTDVELLSHGANSVYGLQSVYPTIIMAGRGHADIYHLLQNGAKYNSDKY 170
; 387 ACKNRKIKIVLELLYHAIETAT-TEGSLPLHVAAGFAGINITYLLQOQANDVAVR 445
; 171 GTPLVMAARKHLECVKHLAMGADVDEGANSTALIVAKGYTSYKELKRNPNV 230
; 446 GETPLHARANOTIVRYLVNRNGAQVDAARELOTPLHISRLGNTDVIILLQANAS 505
; 231 NITDKGNALMASKEGHEIYODLLDAGTYVNIIDRSQDYLIGAVRGHVEIYRAL 290
; 506 NATRDYTPHIAKEGEGEYVAILMDHGTCKTLTKGFTPLHAAKYNGULPYAKSL 565
; 291 QRYADIDIRGQDNKTAIYAVEKGNATMVRDILQCNPDTEICTKDETPLIKATKRNIE 350
; 566 ERGTPEVDEGKNQVPTPLHVAHAYNNDKVALLLLENGASNAKNGYTPHIAKKQMD 625
; 351 VVELLDKAKYSAVDKGDPTPLHVAIRGRSRLAELLRN----- 391
; 626 IASTLHRYANANESKAGFTPLHAAQGHHEMALILENGAKGAOARGSLTPMHICA 685
; 392 -----PRDGLYRPNKAGETPNYND-C-SHOKSILTOIFGARHLSPTEDGMIG 440
; 686 QEDRVSAVELYKENAIDPKTKAGTTPHVAHCH-----FG-----QINMYR 727
; 441 YDLYSSALADILSEPTMOPICVGLYAWGSGKSFLLKL-----EDMKTPAGOTPEL 495
; 728 FIEHGARVSVITRASYP-----LHOAAQOQHNSVRYLLEHGASPNVHTSTG----- 777
; 496 PQFSWLIVPLTLLCGGLVFAFPVDVNTLAIAISLALIVYFYIYFGRGREGESW 555
; 778 -----TPLSI----- 782

```



```

Db 506 NAATRDLYPPLHIAKEGEVEAAIIMDGTDKTLTKKGFPLHIAAKGNLPVAKSL 565
QY 291 OKXADIDIRGODKKTALYMAVEKGNATMVDILQCPDIEICRKOEDPLIKATKRNIE 350
Db 566 ERGTPTVIDSKNOYTPPLVAHAHYNDKVALLLLENGASAAHAAKNQYTPPLHIAAKNMD 625
QY 351 VVELLDKGVASVADKGDTPPLVAIRGRSRIAEILLRN----- 391
Db 626 IASTLLHYNANNAESAGFTPLHIAQESHREMAILLIENGAKVGAQARNGILTPMHLCA 665
QY 392 -----PKGRILYRNKAGETPVNIDCSHOKSILTOIFGARHLSPTETDGMIG 440
Db 686 QEDRVSAEELVKEENAIIDPKTAGYPLHIAACH-----FG-----QINMYR 727
QY 441 YDLSSALADLSEPTMOPPICGLVAGMSSGSKFLKLT-----EDEMKTAGOOTEPL 495
Db 728 FLIEHGKARVSVITRASTP-----LHOAAOQHNSVRYLLEHGASPNVHTSGO----- 777
QY 496 FOFSMILVPLTLICGLGLVFAFPVDTNLAISLFLALYIFFIYIFGGRRGESW 555
Db 778 -----TPLSI----- 782
QY 556 NMAALSTRALHIGIELLFKLMFVNPPPLPBOJTKALPVR-----LFTDYN 604
Db 783 -----AERLGYSVEALKITETTVITETTVTERYKRPONPEAMNETMFS- 830
QY 605 RLSSVGGETSLAEMIALDACCEREGL-ATRLPVRPTEESQCK----- 650
Db 831 --SEDEGEDNOITANNAHAFSESLETKGLHDSIGVHLIHAETPLRSPEVEGTGDDLA 888
QY 651 -----KMKKTCCLPSFVILF-----VCIITAGIT 676
Db 889 LIRKAQHEPITTAADPSLDASLPDNTIMRTMOPSLFSFWDARGAMRCGRISGV 948
QY 677 LLAIFRVPKHLVYNAIILISIASVGLAVLNCRTMOWDLSLN-----SQRRLHSAAS 732
Db 949 IIPPRKAPQTRV-----TCR-YLGDKLHAPPLSEGBAL--AS 986
QY 733 KIKHLKSEG--FMVKLCEVELMARMARKTIDSPONTIRLVII--DG-----LDAC 780
Db 987 RILEMAHGAKEFLQPLVLEVPHPASL-----KGRERIVILRSDGOMKEHOLEAT 1038
QY 781 EODKVLQMLD-----TVRLFSGKPFALIASDPHILIKANLMS 822
Db 1039 E-DAVQEVLENSPDAEELSOLDLHISIRILINDPEM--YFA-----VTVRQEVHC 1090
QY 823 VLKDSNINGDHYKNIV-HLPVPLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTRR 881
Db 1091 VGPBGV-----ILSSVPHVOAIFPDGSLTKTKVVOA-----QP 1127
QY 882 VSQNSLQEMTKL-GSKTALNRDITY--RRROMQRTITROKSPDLTK---LLVTEDMWSDI 935
Db 1128 VPOET---VTRLHGNRAVSPVIVVEPRRKRFRPILCLPLPOSSKGMLTQYSGOPGO 1184
QY 936 SPQTMRLRLIV-----SVYGRLLRANQITFNMRLASWINTLEQ-W----- 976
Db 1185 EPTLRLILCKTGSSSPAQWEDITG---TTQALFTGEDVSTFTTVSARFWIMDCQPRD 1240
QY 977 -----PRTSMILLIYEET--BG-----LPDQMLTKM--YERISKNI 1010
Db 1241 AARMAOEYVNEAIVAPMAKFLI-FARRTPAEQOLRLFCMTDREDKTELEKOFTE-I 1298
QY 1011 PTTKDVPEPL-----LEIDDIRNFVFLSSRPVLVARDVK--FLPCVFN--LDPKL 1059
Db 1299 AKSKDVEVLSGRHOFLFSGNL-----LPTYSQQLSLYFLPEQENRLAEVWKI 1348
QY 1060 R-----EIIADVRAEQINIGLAVPLPLHESPPRPSGQSPASVCSA-----SFNG 1110
Db 1349 RTHDNTADAGR-----IVFMKEPKLAENLPQTP-----VCLAITLPRYTG 1393
QY 1111 PFGGVVSPDPHSSYSGLS-----GPOHPFYNRAAVPATGSSLSLS-----MT 1155
Db 1394 PEP--MWSK---LFYEASILTEKYVGAFHETAEPRDMLPLAHVALLIGADMHRLALEV 1448

```

```

QY 1156 VDVCEKLR-QIEGLDQNMMPQYCTTIKANINGRVLSQCNIDELKEMANFGDMHLR 1214
Db 1449 PIDIRQVRRHQVLGE-----AVTILRIW-----FLKKQALPVA-----LR 1486
QY 1215 SVALEKRVESQVVEPDRF--LNENSSAPVPH--GESARRSS-----HTELPL 1259
Db 1487 SAL--QRIGRDDVREMBRAEKLDELGPVSHISGPSITLSSTLLEVAGDRRRRAEVYM 1544
QY 1260 TELS-SQPYTLNFSF-----ELNTIGLDE-----GAPHSNLSMSO 1297
Db 1545 AOCORLAQEPFOOVYNGTGPDPPEPKDQSFHEEEVAVSIRTVTERHVHDSNGP 1604
QY 1298 TRRTPSLSLNSQDSIEISKLTDY-----QAEY--RDAYREYIAQMSQLEGSTG 1346
Db 1605 IVEERTITTTEDDAVANEELIVDKIVPLNEEQKEMBRMAREVMEFQOETSKEGTFG 1664
QY 1347 SSTIGRSSPHSTYYIGOSSGGSJHSTLEORGRGEGELKODGRKSLMRGVDYSS 1406
Db 1665 CQT-----THEKEDDGSLSKTTW-----KDSHVQ-----IFPDGGETSANE 1703
QY 1407 SGVSTNEASP-LDPTIETD 1424
Db 1704 TGLSSGADTMTPTTKED 1722

```

## RESULT 5

```

US-09-065-474-33
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIPOPTILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-065-474-33

```

```

Query Match 5.3%; Score 466.5; DB 3; Length 1745;
Best Local Similarity 20.4%; Pred. No. 1e-23;
Matches 358; Conservative 235; Mismatches 575; Indels 591; Gaps 73;

```



APPLICATION NUMBER: US/09/557,034  
 FILING DATE: 21-Apr-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/065,474  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Verser, Carol Talkington  
 REGISTRATION NUMBER: 37,459  
 REFERENCE/DOCKET NUMBER: HR-5-C1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 970/493-7272  
 TELEFAX: 970/484-9505  
 INFORMATION FOR SEQ ID NO: 33:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1745 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
 US-09-557-034-33

Query Match 5.38; Score 466.5; DB 4; Length 1745;  
 Best Local Similarity 20.4%; Pred. No. 1e-29;  
 Matches 358; Conservative 235; Mismatches 575; Indels 591; Gaps 73;

14 EENIPALKALECKVDNERNGGOTPLLAEOGVVEIVKELINKGANCLEDIDN-- 71  
 Db KDDTKATLIDLOENHSIDVTSKSGFTPLHAHGENVAQILLLEGAVANVQARHNS 266  
 Qy 72 -----W-----TALISAKKEGHIHYEELIKSGASD--- 97  
 Db 267 PLHATKMGRTNNVSLLAGAVIDCRTLPLPHCASRSGHDYVDLLEKGPISAK 326  
 Qy 98 -----W-----EHRDMGGWALMW 110  
 Db 327 TKNGLAPLHMAQYDVTVDYLPPLHVAHCGHVRAKLLDRNAPNARALNGFTPLHI 386  
 Qy 111 ACYKGTVDYVELLSHGANSVTLGQSVYPIIMAGRGHADVHILLONGAVNCSDKY 170  
 Db 387 ACKNRKIKYELKTKHAAIEAT-TEGSLSPHVAAMGAINIVTYLLQOGAADAATYR 445  
 Qy 171 GTPPLVAARKGHLECYKHLLAGADYDQGANSMALIVAAGGYTQSYKEILKRPNV 230  
 Db 446 GEPPLHAAARANGDIYRLVBRGAQYDAABELQPLHSHSLGNDIYILLQANASP 505  
 Qy 231 NLDDKGNLTALMAASKGHEIYVODLDACTYVNIIDRSGDYVLIGAVRGHAYEIRAL 290  
 Db 506 NAATRDLYPLHIAAKGGEVAAILMDHGTDLTKLTKKGTPLHIAAKYGNLPVAKSL 565  
 Qy 291 OKYADIDIRGODNKTALYMAVEKGNATVMDILOCNPDEICTKDGFTPLIKATKMANTE 350  
 Db 566 ERGPVIVIEGKNQVTPPLHVAHYNNDKVALLLLENGASAAAKNGTTPHIAAKKNQMD 625  
 Qy 351 VVELLDKGAKVASVDKGDTPPLHVAIRGRSRLAEILLRN----- 391  
 Db 626 IASTLLHXYANANASAKGFTPLHIAOEGHREMAALLIENGAKVGAQANGLTPHILCA 685  
 Qy 392 -----PKDGLILRPKNAGETPYNIIDCSHOKSILTQIFGARHLSPTEDGMLG 440  
 Db 686 QEDRVSAEELVVENAIDPKTAGYTPHVAACH-----FG-----QINMVR 727  
 Qy 441 YDLSALADILSEPTQPIICVGLVQWSSGSKFLKTL-----EDEMKTFAQOQTEPL 495  
 Db 728 FLIEHGRVSVITRASTP-----LHQAOOGHNSVRYLLEHGASPNVHTSTGO----- 777  
 Qy 496 FQESWLVPLTLICGLIGLVFAFPVVDNLAIASISFLAIYIFIVIFGGRREGESW 555  
 Db 778 -----TPIST----- 782  
 Qy 556 NMAALSTRARHIGYELLFKLMFVNPPPELPEOTTKALVRF-----LFTDYN 604  
 Db 783 -----AERLGYSVVEALKTITETVITETTVTEREYKPPONPEAMNETMFSD-- 830

Qy 605 RLSSVGENSLAEMLATLSDACEREGFL--ATRLFVRVREESGKK----- 650  
 Db 831 --SEDEGDNQITANNAHADSESLTKGHDSTGVNLIHAEPTLSRSEVEEGIDGDDA 888  
 Qy 651 -----NMKTCCLSPVIFLEFI-----VGLIACIT 676  
 Db 889 LIRKAQHEDITAMADPSLADSLPNDVITMRTMQPSFLSPMDVARGAMRGCHSGVR 948  
 Qy 677 LAIFRVDPKHLTVNALISIASVGLAFVLCRTWQVYLSLN-----SQRKRHSAS 732  
 Db 949 IIPPRKAPQPTRV-----TCR--YLGKCKLAHPPLSGEAL--AS 986  
 Qy 733 KLHLKLSSEG--FMKYLKCEVELAMAMAKTIDSFQNOTRLVYII--DG-----LDAC 780  
 Db 987 RILEMAPHGAKFLGVILFVPFASL-----KGREREIYILSDGQMKHQLEAT 1038  
 Qy 781 EODKYLQMLD-----TVRLFSKGFPIALFASDPHIIKAINQNLNS 822  
 Db 1039 E-DAQVEVLNESFDAEELSOLDLHTSRITRLTNDPFW--YFA-----VTVRQOEYHC 1090  
 Qy 823 VLKDSNINHDMRNIV-HLPVFLNSRGLSNARKLYSATNGDITCSDTGTGDEPTRR 881  
 Db 1091 VQPEGV-----ILSSVPHVQALFPDGSJTKTKYVQA-----QP 1127  
 Qy 882 VSQNSLGEWTKL-GSKTALNRDQY--RRQKQRTITROMSPDLTK--LVTEDMFSDI 935  
 Db 1128 VFOET--VTRLHGRNVAASPIYVIEPRRKRHKITLCIPLPQSSKMGMLTQYSGQFQ 1184  
 Qy 936 SPQTMRLNIV-----SVTGRLLRANOITENMRSLASMINLQD-W----- 976  
 Db 1185 EPTTLRLCSKSGSSPAQMEDITG---TQDLFTGEDVSPFTTVASRFLMDCQTPRD 1240  
 Qy 977 -----PYKSWILYIEE--EG-----LPDQTLTKM--YERISKNI 1010  
 Db 1241 AARMAQEVYNEAIVAPYAKFLI-FARRFPREGOLRLFCMDDEDEDTLEKEERPI-I 1298  
 Qy 1011 PTKDVEPL-----LEIDGRNEFEVFLSSRTPLVAVDYVT--FLCTVW--LDPKL 1059  
 Db 1299 AKSKDEVLSGRHOFLEFSGNL-----LPTKSSDQSLFTLPQEKRLAFWKI 1348  
 Qy 1060 R-----ELIADYRAAREQINIGLAYPPLHGGPPRPSGYSQAPVYSSA--SENG 1110  
 Db 1349 RTHDNETAADR-----IVFMKEPKLRAENLPPQF-----CTLAITLPEYTG 1393  
 Qy 1111 PRPGVYSPQPHSSYSGIS-----GPHPRYNAAPATGSSILLSS-----MT 1155  
 Db 1394 PEP--MYSKR-----LEFSEASLTKYVGAFHETAEEDNPLPLAHVALLIGADWHRLARALEV 1448  
 Qy 1156 VDVCSEKLR-QIEGLDQNMPOYCTTIKKNINGRVLSCNIDELKEKAMFGDMHLFR 1214  
 Db 1449 PIDIDQVHQLVGLF-----AVTILRIW-----FLKKQATPYA--LR 1486  
 Qy 1215 SVALENRVSQVPEDEPRF--LNENSSAPVPH--GESARRSS-----HTELPL 1259  
 Db 1487 SAL--ORIGRDVYREMDRAEKLDGEGPVSHISGPTSTLSTLLEVAGDRRAHAYTM 1544  
 Qy 1260 TELS--SQPYTLNFSF-----ELNTLGLDE--GAPRHSNLSMSQ 1297  
 Db 1545 AAOQLAQQEPFOQVNGTGPDEBPKEQSFHEEEVAVSEIRTVVTERIVHDSNGP 1604  
 Qy 1298 TRTPSLSLNSQDSSIEISKLTDV-----QAEY-----RDAREYIAQMSLEGSTG 1346  
 Db 1605 IYEBKITITTYDDAVANEELIVDKIPLNNEEOKMDRMAVEEMNEQDQTSKEGTFG 1664  
 Qy 1347 SSTISGRSSPHSTYYIGSSSGSISHSTLEOERGEKELKODKRSFLMKRGDVIDYS 1406  
 Db 1665 CQT-----THEKEDDQGSILKTTM-----KDSHVRO-----IFPDGGETSANE 1703  
 Qy 1407 SGYSTNEASP-LDPTTEED 1424  
 Db 1704 TGLSGDAQDTIMPTTKED 1722



RESULT 7  
 US-09-082-059-2  
 ; Sequence 2, Application US/09082059A  
 ; Patent No. 6225086  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Devorajan, Jon S.  
 ; APPLICANT: Morrow, Prasad  
 ; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identification  
 ; FILE REFERENCE: 44574-5002-US  
 ; CURRENT APPLICATION NUMBER: US/09/082,059A  
 ; EARLIER FILING DATE: 1998-05-21  
 ; EARLIER APPLICATION NUMBER: 60/047356  
 ; EARLIER FILING DATE: 1997-05-21  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1088  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-082-059-2

Query Match 5.1%; Score 449.5; DB 4; Length 1088;  
 Best Local Similarity 31.1%; Pred. No. 1.2e-28;  
 Matches 118; Conservative 74; Mismatches 184; Indels 3; Gaps 2;

QY 14 EEENIPALKLECKKCDYDERNEGOTPLMLAEGNGVEIKELKNGANCNLEDLDMWT 73  
 DB 30 KNRIRKVELLKHGASIOAYTESGLTPIHYAAEFHGHVITYSOLMHHGASPTTNVNGET 89  
 QY 74 ALISAKSGHIIHVEELLKSGASLEHRDMGWTALMAACYKRTDVEYELLSHGANP-SV 132  
 DB 90 ALHMAARSQAEVNVRYLVODGAQVAKAKDQPTPLHISARLGKADIVQOLLQOGASPMAA 149  
 QY 133 TGLQSVYPIITMAARGHADIYHLLONGAKVNCSDKGTPTPYMAAKGHLECYKHLA 192  
 DB 150 TTSGTT--PLHLSAREGHEDVAALFLDHGASISITTKKGFPLHYAAAYGKLEVANLLQ 207  
 QY 193 MGADYDOEGANSMTALIVAVKGYQSVKEILKRNPNVNLTDKDGNTALMAASKGHEI 252  
 DB 208 KASPDAGKSGLTPLHYAAHNDKQVALLDQASPHAAAKNGYTPPLHIAAKKNQDI 267  
 QY 253 VODLDAGTYNIPRSGDVTYLIGAVRGHVEIYRALLQYADIDIRGDNKKTALYMAVE 312  
 DB 268 ATTLEYGADANAVYRGIASVHLAAQEGHVDVSLIGRANAVLSNKSGLTFLHLAAQ 327  
 QY 313 KGNATMVRDILQCNPDTEICTKDETPLIKATKKNIEVEYELLDDKGAASVADKGDTP 372  
 DB 328 EDRVAVAEVLVNOGAHVDAQTKMGYTPPLHVGCHGNIKIVNELLQHSKAKVNAKTKNGTTP 387  
 QY 373 LHVAVRGRSRRLAEELLRN 391  
 DB 388 LHOAAQGHHTHINVLQN 406

RESULT 8  
 US-08-810-712-10  
 ; Sequence 10, Application US/08810712G  
 ; Patent No. 6160106  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yeda Research and Development Co. Ltd  
 ; TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and  
 ; FILE REFERENCE: sequence list  
 ; CURRENT APPLICATION NUMBER: US/08/810,712G  
 ; CURRENT FILING DATE: 1997-03-03  
 ; EARLIER APPLICATION NUMBER: PCT/US94/11598  
 ; EARLIER FILING DATE: 1994-10-12  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 1423  
 ; TYPE: PRT

ORGANISM: Homo sapiens  
 US-08-810-712-10

Query Match 4.7%; Score 414.5; DB 4; Length 1423;  
 Best Local Similarity 21.6%; Pred. No. 1.9e-25;  
 Matches 221; Conservative 156; Mismatches 335; Indels 309; Gaps 36;

QY 5 ISQSVINVEENIPALKLECK--DVERNEGOTPLMLAEGNGVEIKELKNGA 62  
 DB 336 VMALITHAIDNDVPGLOHLSLSNVDNQPKHTPPLLIAGCGNIQLILTKRGS 395  
 QY 63 NCNLEDNTALISAKSGHIIHVEELLKSGASLEHRDMGWTALMAACYKRTDVEYEL 122  
 DB 396 RIDVQDGSGNAYMAARHGHVDTPLFLSKNKPPLVYKSGEMALHVARAGHADVAV 455  
 QY 123 LSHGANPSTVGTQYVPIITMAARGHADIYHLLONGAKVNCSDKGTPTPYMAARG 182  
 DB 456 TCASAOIPISRTKEETPLHCAAMHGYSAKALCEACGNVINKRBEETPLTPASARG 515  
 QY 183 HLECYKHLAMGADVDOEGANSMTALIVAKGYQSVKEILKRNPNVNLTDKDGNTALM 242  
 DB 516 YHDIVECLAHGAD-----LACDKGHIALH 542  
 QY 243 IASKEGHEIYVODLDAGTYNIPRSGDVTYLIGAVRGHVEIYRALLQKYADIDIRGOD 302  
 DB 543 LAVRCOMEVITKLLSGCFVDYQDRHNGTPLHVACKDGNMPTVALCE----- 591  
 QY 303 NKTALVAVKGNATVYRDLQCNPDTEICTKDETPLIKATKKNIEVEYELLDDKGA 362  
 DB 592 -----ANCNLD--ISNKGRTPLHAAANGILDVRYLCIMGASV 629  
 QY 363 SAVDKGDTPLHYAIRGRSRRLAEELLRNPKDRLYRNKAGETPNIDCSHOKSILQ 422  
 DB 630 EALTDTCKTMDLARBEGHEHVAGLARLKD-----THRGFLIQ 670  
 QY 423 IFGARHLSPTETDGMIGDYSSALADILSEPMPPICVGLYQAQSGKSFLLKLED 482  
 DB 671 -----LRPTQ-----NLQPRIKLKFHSGSGKTTIVESTK- 701  
 QY 483 EMKTFAGQGTPELPQSWMLVFLTLGLGLGVF-----AFPDNTNA 526  
 DB 702 -----CGLNSFFRRRRPRSSINSSRPPSPPLASKET 734  
 QY 527 IATSLSEFLALITYFFVIYFGRRBESNMAMALSTRLARHIGYLELLFKLFPVNPRL 586  
 DB 735 VSVSINN-----YPCGENSVSRSMPEPGLK--GMLE-----VFAPRTH 776  
 QY 587 P-----EQTTAL-----PYRPLFTDYNRLSVSGETISLAEMTA 620  
 DB 777 PHCSADDOSTKAIDIONAYLVNGVDFSVWFSGNPPYFCCYDY--FAANDPTSIHVYV 833  
 QY 621 TLDACERERF--FLATRLFRVFRTEES--QGKKMKKTCCLPSEFLFYVGLIING 674  
 DB 834 SLEPEYEQDNPVITFLWSFLKSLVPEEPLRAGGKL-----NPLQVVL----- 877  
 QY 675 ITLLAIFRVPKHLTVNALISASVYGLAFVLCNRTWQVLDLSLNSQKR--LHSA 730  
 DB 878 -----VAHADIIMVPRPAGGEFGDXKT-----SLKTEIRNFGDLH-I 917  
 QY 731 ASKLHLKSGEFG--MKVLKCEVELAMAMARTIDSTFONQRLVYIIDGLDACEQDKYL 786  
 DB 918 SNKLFVLDGASGSKDKVLRNHLQ--EIRSOIVSVCPPMHL-----CE--KII 963  
 QY 787 QMDTVRVLESKGPFAI--FASDPHIIIKAINONNSLRDINSINGHDMRYNIHLPEV 844  
 DB 964 STLPWRKLNGPQLMSLDQFYD-----VODQINPLASED-----LRRIAQ-- 1007  
 QY 845 LNSRGLSNARKFLVTSATNGDITGSDPTGTOEDTRVYSONSLGENTKIGSKTALNR-RD 903  
 DB 1008 LHSSTGEIN--IMSETYQVDVLLD-----PWLCTNVLGKLLSTEFRLAHNRG 1055  
 QY 904 TYRRQMQRTITRQMSFDITKLLVTEDMFS-DISPTMRRLINIVSVTGRLIRANOITFN 962

Db 1056 RTVEDIORTLVDSVDEELQILDAMDICARDLSSGTM-----VDYPA-LIKTDNLHRS 1108  
QY 963 W 963  
Db 1109 W 1109

RESULT 9  
US-09-196-387-2

Sequence 2, Application US/09196387  
Patent No. 6277613  
GENERAL INFORMATION:  
APPLICANT: de Lange, Titia  
APPLICANT: Smith, Susan  
TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
TITLE OF INVENTION: OF USE THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/196,387  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/095,225  
FILING DATE: June 10, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1327 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-09-196-387-2

Query Match 4.5%; Score 395.5; DB 4; Length 1327;  
Best Local Similarity 23.3%; Pred. No. 6,9e-24;  
Matches 167; Conservative 80; Mismatches 211; Indels 259; Gaps 17;

QY 14 EENIRPAKALKKCKDVERNECGOTPLMLAEGQVETVELKNGANCNLEDLDMNT 73  
Db 348 EERIMLLTPPLNCHASDGRK---STPLHLAAGYNKRVIVOLLQHGADVHAKDKGGLV 404  
QY 74 ALISASKEGHIHVEELKSGASLEHRDMGWTALMACYKGTDTVELLSHGANSV- 132  
Db 405 PLINACSYGHEYTELLKKGACVNAADMQLPPLHBAASKNVEVCSLLSHGADPTLV 464  
QY 133 -----T 133  
Db 465 NCHGSAVDMAPTPELRELTVEFKGSHLQAAREADLAKVKTALLETINRQPSHET 524  
QY 134 GLQSY-----YPIIWAAGRHADIVHLLONGARV 164  
Db 525 ALHCAVASLHPKRYOTELLRKGANVNEKNDFMTPLHVAARAHNDVMEVLHKHGAKM 584

QY 165 NCSDRYGTPPLVMAARKHLECYKHLILANGADVDGANSMTA----- 207  
Db 585 NALDTLGGTALHRAALAGHLQTCRLLISGSPSISLGGFRAQNGNVAQOILSESP 644  
QY 208 -----LIVAVKGYTQSVKEILLKRNPNVLTDKDG--NTALMIASKEGHIETVDL 256  
Db 645 IRTSDVDYRLLESKAGDLETVKQLCS--SQNVNCRDLERHSTPLHFAGYNRVSVEXL 703  
QY 257 LDAGTYVNIPIRSGDVLIGAVNGHEIVYRALLQKRYADIRGQDNKALYAVEKG-- 314  
Db 704 LHHGADVHAKDKGGLVPLINACSYGHEYVAELLVRGASVNVADMLKFTPLHBAAKGY 763  
QY 315 -----NATVRDILQ-----C 325  
Db 764 EICKLLKHGADPTKKNRDNPNPLDVKEGDPDIQDLKGDALLDAAKGGLARYOKC 823  
QY 326 NPDTLEICTD---GETPLIKATKRNIEVEVLLDKGAVSAVDKKGTPLH----- 374  
Db 824 TPEINCRDTQGRNSTPLHLAAGYNLVEAEYLLERGADVNADKGGLPLHNAASYGHV 883  
QY 375 -----VAIRGSRRLAEILLRNPDKGRLLYPRNKAETP 408  
Db 884 DIAALLIKYNTCVNATDKKATFPLHBAOKGRT--QICALLHAGADPTM--KNQGGTP 939  
QY 409 YN-----IDCSHOKSILT-----QIFGARHLSPETDGMGLDYSSALADI 451  
Db 940 LDATADDIRALLIDAMPPEALPTCFQATVVASLISPASTPS-----CLSAASISDN 994  
QY 452 LSEPTMQPPICVGLVQWQSG-----KSFLLKLEDEKMT 486  
Db 995 LIGPLAE--LAVGASNAGDGAAGTERKEGEVAGLDMNISQPLKSLGLEHLNDIET 1049

RESULT 10  
US-09-196-387-10

Sequence 10, Application US/09196387  
Patent No. 6277613  
GENERAL INFORMATION:  
APPLICANT: de Lange, Titia  
APPLICANT: Smith, Susan  
TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
TITLE OF INVENTION: OF USE THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/196,387  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/095,225  
FILING DATE: June 10, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:

LENGTH: 949 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-196-387-10

Query Match 4.4%; Score 387.5; DB 4; Length 949;  
 Best Local Similarity 25.2%; Pred. No. 1.8e-23;  
 Matches 144; Conservative 59; Mismatches 177; Indels 191; Gaps 11;

14 EENIPALKALLEKCKVDNERNEGGQPLMLAAGVNEIVKELLKNGANCNEEDDNT 73  
 348 EKKMLALPLNVNCHASDGRK---STPLHLAAGYNNRYIYOLLHGADVHAHAKGGLV 404  
 74 ALISAKSGHITHYELLKSGASLEHNDMGSTLMMACKYGRDVEYELLSHGANSV- 132  
 405 PLHNACSTGYEVELLKHGACVNMADLMOWFTPLHBAASKNRVECVSLSHGADPTLV 464  
 133 -----T 133  
 465 NCHGSAVDMAPTPELRERLTERKSHSLQAAREADLAKYKTLALEIINFKQPSHET 524  
 134 GLQTSV-----YPLIWAAGHGADYHLLQNGAKV 164  
 525 ALHGAVALSLHPRKQVTELLRKGANVNEKNKDMFTPLHVAERAHNDVMEVLLKHGAKM 584  
 165 NCSKYGTPPLVMAARKGHECVKHLAMGADVDEGANSMTA----- 207  
 565 NALDPLGOTALLHRAALAGHLOTCRLLSYSGSPSTISLQGTFAOMGNEAVOQLSSTP 644  
 208 -----LIIVAGGYTOSVKEILKRPVNLTRDKG--NTALMTASKGHEIYVDL 256  
 645 IRTSDVDRLEASAGDETVKDCS--SQVNCRDLEGRHSTPLHFAAGYNNVSYEYL 703  
 257 LDAGTYVPIPRSGDTVLIGAVRGHVEIVALLQKADIDIRQDNKTALYMAVEG-- 314  
 704 LHHGADVAKDKGLVPLHNACSYGHEVAELVHHGASVNAVLMKFTPLHFAAAGKY 763  
 315 -----NATVVDILQ-----C 325  
 764 EICKLLKHGADPTTKKNDGNTPLDYKESDPTDLODLKGAALLDAKKGLARVOKLC 823  
 326 NPDTEICTKD---GETPLIKATKRNIEVEVELLDKGAVSADVKKGDTPLHVAIRGRSR 382  
 824 TPEINCRDFOGRNSTPLHLAAGYNNLEVAEYLLHGGADVNAODKGLIPLHNA----- 877  
 383 RLAEILLRPNKDGRLVPRNKAGETPVNIDC 413  
 878 -----ASYGGLARVOKLC--TPEINIC 898

US-09-031-485-28  
 Sequence 28, Application US/09031485  
 Patent No. 5824306

GENERAL INFORMATION:  
 APPLICANT: Tang, Liang  
 APPLICANT: Blehm, E. Scott  
 TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
 TITLE OF INVENTION: USES THEREOF  
 NUMBER OF SEQUENCES: 85  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Carol Talkington Verser, Ph.D.  
 ADDRESSEE: Heska Corporation  
 STREET: 1825 Sharp Point Drive  
 CITY: Fort Collins  
 STATE: Colorado  
 COUNTRY: USA  
 ZIP: 80525  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows 95  
 SOFTWARE: WordPerfect for Windows, Version 7.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/031,485  
 FILING DATE:  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/847,429  
 FILING DATE: 24-APR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Verser, Carol Talkington  
 REGISTRATION NUMBER: 37,459  
 REFERENCE/DOCKET NUMBER: HW-5  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 970/493-7272  
 TELEFAX: 970/484-9505  
 INFORMATION FOR SEQ. ID NO: 28:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 348 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-031-485-28

Query Match 4.2%; Score 376; DB 2; Length 348;  
 Best Local Similarity 34.0%; Pred. No. 2.5e-23;  
 Matches 113; Conservative 47; Mismatches 134; Indels 38; Gaps 4;

45 AAGQVNEIVKELLKNGANCNEEDDNTMTALISASKEGHIHYVELLSGASLEHRDNG 104  
 44 AARAGNDRLVELLEKSGDTINTCNANGALHLASKEGHEVYRELKRAVDATKRG 103  
 105 WTALMACYKGRDVEVELLSHGANSPTVGLQSYVPIIWAAGRGHADIHLLONGAKV 164  
 104 NTLHLASLAGDELITVLEVGNANVQSLN--GFTPLMAAGNEHSESVRYLHANNO 162  
 165 NCSKDTGTPPLVMAARKGHECVKHLAMGADVDEGANSMTALYVAGGYTOSVKEIL 224  
 163 ALSTEDGFTPLVALAOGGHDVAVALL---ENDTRGVKRLPALHLAAKDDTKATLL 218  
 225 KRPNVNLDPKDNFTALMTASKGHEIYVDLIDAGTYVPIPRSGDTVLIGAVRGHVE 284  
 219 QNEHNSDVTSKSGFTPLHLAAGYNNVQALLERKANVNAQARHNSISPLHATKGRIN 278  
 285 IYVALLQKADIDIRQDNKTALYMAVEGNATVVDILQCPDTEICTKDEPTPLIKAT 344  
 279 MVSLLAHGAVIDCR-----TRDLL-----TFLHCAS 305  
 345 KMRNIEVELLDKGAKVSADVKKGDTPLHVA 376  
 306 RSGHDQVVDLLERKGAIPISAKTKNGIAPLHMA 337

US-08-847-429A-28  
 Sequence 28, Application US/08847429A  
 Patent No. 5827692

GENERAL INFORMATION:  
 APPLICANT: Tang, Liang  
 APPLICANT: Blehm, E. Scott  
 TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
 TITLE OF INVENTION: USES THEREOF  
 NUMBER OF SEQUENCES: 85  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Carol Talkington Verser, Ph.D.  
 ADDRESSEE: Heska Corporation  
 STREET: 1825 Sharp Point Drive  
 CITY: Fort Collins  
 STATE: Colorado  
 COUNTRY: USA  
 ZIP: 80525  
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/847.429A  
FILING DATE: 24-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-847-429A-28

Query Match 4.2%; Score 376; DB 2; Length 348;  
Best Local Similarity 34.0%; Pred. No. 2.5e-23;  
Matches 113; Conservative 47; Mismatches 134; Indels 38; Gaps 4;

QY 45 AAEQGNVEIYKELKNGANCLELDNMTALISASKEGHIIVEELLKSGASLEHRDMG 104  
DB 44 AARGNDRIYELRLSGDITCNANGLNALHLASKEGHEVRELLKRAKDVAATRG 103  
QY 105 WTALMAYCYGRDVEYVLLSHGANPSYGLQSVPIIWAAGRGADIVHLLONGAKV 164  
DB 104 NTAHLASLAGOEIIVLYVNGANVVOGLN-GFTPLVMAQENHESVYRLLANANQ 162  
QY 165 NCSKDYGTPLVMAARKGHECYKHLAMGADVDEGANSMTALIVAGGYTOSVKEIL 224  
DB 163 ALSTEDGTFPLVALQGHDRVAVLL-----ENDTRGKVRRLPAHLIAKKDDTKAATLL 218  
QY 225 KRPNVNLTDKDNLTALMASKEGHEIYVODLDAGTYVNIIPDSGDTVLIGAVRGHVE 284  
DB 219 QNEHNSDVTSKSGFTPLHIAHNGENVAQLLEKGANVYQARHNISPLHATKMGRTN 278  
QY 285 IYRALLQYADIDIRGQDNKTALYMAVEKGNATVNDIQCNDPEICTCKDDETLIKAT 344  
DB 279 MVSLLAHGAVIDCR-----TRDL-----TPLHCAS 305  
QY 345 KMRNIEVELLDKGAKVSAVDKKGDTPLHVA 376  
DB 306 RSGHDQYVDLLEKGAPEISAKTRKNGIAPLHMA 337

RESULT 13  
US-09-065-474-28  
Sequence 28, Application US/09065474  
Patent No. 6063599  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
NUMBER OF SEQUENCES: 171  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/065.474  
FILING DATE: 24-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5-C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-065-474-28

Query Match 4.2%; Score 376; DB 3; Length 348;  
Best Local Similarity 34.0%; Pred. No. 2.5e-23;  
Matches 113; Conservative 47; Mismatches 134; Indels 38; Gaps 4;

QY 45 AAEQGNVEIYKELKNGANCLELDNMTALISASKEGHIIVEELLKSGASLEHRDMG 104  
DB 44 AARGNDRIYELRLSGDITCNANGLNALHLASKEGHEVRELLKRAKDVAATRG 103  
QY 105 WTALMAYCYGRDVEYVLLSHGANPSYGLQSVPIIWAAGRGADIVHLLONGAKV 164  
DB 104 NTAHLASLAGOEIIVLYVNGANVVOGLN-GFTPLVMAQENHESVYRLLANANQ 162  
QY 165 NCSKDYGTPLVMAARKGHECYKHLAMGADVDEGANSMTALIVAGGYTOSVKEIL 224  
DB 163 ALSTEDGTFPLVALQGHDRVAVLL-----ENDTRGKVRRLPAHLIAKKDDTKAATLL 218  
QY 225 KRPNVNLTDKDNLTALMASKEGHEIYVODLDAGTYVNIIPDSGDTVLIGAVRGHVE 284  
DB 219 QNEHNSDVTSKSGFTPLHIAHNGENVAQLLEKGANVYQARHNISPLHATKMGRTN 278  
QY 285 IYRALLQYADIDIRGQDNKTALYMAVEKGNATVNDIQCNDPEICTCKDDETLIKAT 344  
DB 279 MVSLLAHGAVIDCR-----TRDL-----TPLHCAS 305  
QY 345 KMRNIEVELLDKGAKVSAVDKKGDTPLHVA 376  
DB 306 RSGHDQYVDLLEKGAPEISAKTRKNGIAPLHMA 337

RESULT 14  
US-09-557-034-28  
Sequence 28, Application US/09557034  
Patent No. 6365569  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
NUMBER OF SEQUENCES: 171  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

```
Search completed: July 1, 2003, 14:40:44
Job time : 29.5 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 89 Seconds

(without alignments)  
3970.458 Million cell updates/sec

Title: US-10-021-571-4

Perfect score: 8884  
Sequence: 1 MSYLISQSYINVEENIPALAKLEKCKVDNERECGOTPLMAEGNLTIVELIKN 1715

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_plant:\*  
10: sp\_rodent:\*  
11: sp\_virus:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteria:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	8830	99.4	1777	4	Q9ULH0
2	8362.5	94.1	1762	11	Q9EGG6
3	8266	93.0	1715	11	Q9ERD4
4	5205.5	58.6	1031	4	Q9UF42
5	2791	31.4	543	4	Q9H9E4
6	2379.5	26.8	1498	5	Q9W211
7	1832.5	20.6	1398	5	Q20109
8	1255	14.1	246	4	Q9NT37
9	666	7.5	129	4	Q9H889
10	557.5	6.3	1856	4	Q99407
11	544	6.1	426	16	Q8YTG9
12	523	5.9	1943	11	Q61307
13	522.5	5.9	1726	11	Q8VC68
14	522	5.9	1719	4	Q13768
15	521	5.9	4377	4	Q12955
16	519.5	5.8	2443	5	Q9VSA2

17	517	5.8	2622	11	Q70511	070511	ratus norv
18	506	5.7	1848	11	Q61302	061302	mus musculu
19	500.5	5.6	6894	5	Q17343	Q17343	caenorhabdi
20	498.5	5.6	1136	6	Q9N180	Q9N180	bos taurus
21	495.5	5.6	1159	5	Q9NCP8	Q9NCP8	drosophila
22	493.5	5.6	843	11	P97582	P97582	ratus norv
23	488.5	5.5	1486	4	Q8REF1	Q8REF1	homo sapien
24	485.5	5.5	1009	5	Q8SWY2	Q8SWY2	drosophila
25	482.5	5.4	784	5	Q9YCA7	Q9YCA7	drosophila
26	482.5	5.4	4001	5	Q8MR07	Q8MR07	drosophila
27	481.5	5.4	833	4	Q96186	Q96186	homo sapien
28	478.5	5.4	1549	5	Q9Y4B1	Q9Y4B1	drosophila
29	475.5	5.4	1549	5	Q24241	Q24241	drosophila
30	474.5	5.3	1599	11	Q99NH0	Q99NH0	mus musculu
31	474	5.3	1762	11	Q88521	Q88521	ratus norv
32	466.5	5.3	1188	4	Q9H288	Q9H288	homo sapien
33	463	5.2	1282	5	Q8T4F8	Q8T4F8	drosophila
34	463	5.2	2119	5	Q9YAU5	Q9YAU5	drosophila
35	463	5.2	2119	5	Q8T9G9	Q8T9G9	drosophila
36	461	5.1	1867	5	Q17486	Q17486	caenorhabdi
37	447.5	5.0	1815	5	Q17488	Q17488	caenorhabdi
38	447	5.0	460	4	Q9H6J9	Q9H6J9	homo sapien
39	447	5.0	743	4	Q8T846	Q8T846	homo sapien
40	442.5	5.0	1809	5	Q17487	Q17487	caenorhabdi
41	442	5.0	2039	5	Q17489	Q17489	caenorhabdi
42	441	5.0	627	4	Q96G77	Q96G77	homo sapien
43	441	4.9	686	4	Q9NXF0	Q9NXF0	homo sapien
44	438	4.9	1088	4	Q13484	Q13484	homo sapien
45	437	4.9	532	4	Q9H2U0	Q9H2U0	homo sapien

## ALIGNMENTS

RESULT 1  
Q9ULH0 PRELIMINARY; PRT; 1777 AA.  
ID Q9ULH0  
AC Q9ULH0;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE K1A1250 protein (Fragment).  
GN K1A1250.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=20039619; PubMed=10574462;  
RA Nagase T., Ishikawa K., Kikuno R., Hirotsawa M., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XV.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 6:337-345(1999).  
DR EMBL: AB033076; BAA86564.2; -.  
DR InterPro: IPR002110; ANK.  
DR Pfam: PF00023; ANK; 12.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00248; ANK; 11.  
DR PROSITE: PS50088; ANK\_REPEAT; 10.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
KW ANK repeat; Repeat.  
FT NON\_TER  
SQ SEQUENCE 1777 AA; 197209 MW; B6505923FBA5F143 CRC64;  
Query Match 99.4%; Score 8830; DB 4; Length 1777;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 1714; Conservative 0; Mismatches 0; Indels 58; Gaps 2;  
1 MSYLISQSYINVEENIPALAKLEKCKVDNERECGOTPLMAEGNLTIVELIKN 60  
|||||

Db 7 NSVLSQSVINVEEENTPALKLEKCKDVERNECCOTPLMIAEONGEIVELIKN 66  
 QY 61 GANCLIEDNDNTALISASKEGHVHIVEELLKGVNLEHRDMGWTALMAMCYKRTD 120  
 Db 67 GANCLIEDNDNTALISASKEGHVHIVEELLKGVNLEHRDMGWTALMAMCYKRTD 126  
 QY 121 ELLLSHGANSVTVGLTISVYPIIMAGRGHADVLLQNGAKVCSKYGTPPLVMAAR 180  
 Db 127 ELLLSHGANSVTVGLTISVYPIIMAGRGHADVLLQNGAKVCSKYGTPPLVMAAR 185  
 QY 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGYTOSVKEILKRNPNVLLDKGNTA 240  
 Db 186 KGHLECVKHLAMGADVDOEGANSMTALIVAKGYTOSVKEILKRNPNVLLDKGNTA 245  
 QY 241 LMIASKEGHTIVOLLNAGTYVNPDRSGDVLIGAVRGHVEYVRLLOKRYADIDIRG 300  
 Db 246 LMIASKEGHTIVOLLNAGTYVNPDRSGDVLIGAVRGHVEYVRLLOKRYADIDIRG 305  
 QY 301 QDNKTALYWAVERGNATWVRLDLOCNPDEICTKGETPLIKATKRNIEVELLDKGA 360  
 Db 306 QDNKTALYWAVERGNATWVRLDLOCNPDEICTKGETPLIKATKRNIEVELLDKGA 365  
 QY 361 KVSAYDKGDPPLHAIAGRSRKLAEELLRNPKDGRLLYRNKAGETPNIDCSHOKSIL 420  
 Db 366 KVSAYDKGDPPLHAIAGRSRKLAEELLRNPKDGRLLYRNKAGETPNIDCSHOKSIL 425  
 QY 421 TOIFGARHLSPETEGMDIGYDLSALADILSEPTMPPICVGLYAAWGSKSTLKL 480  
 Db 426 TOIFGARHLSPETEGMDIGYDLSALADILSEPTMPPICVGLYAAWGSKSTLKL 485  
 QY 481 EDEMTFAGOOIEPLFOFSWMLVFLTLGGLGGLFAFYVHPNGLIVSEFLALYIF 540  
 Db 486 EDEMTFAGOOIEPLFOFSWMLVFLTLGGLGGLFAFYVHPNGLIVSEFLALYIF 545  
 QY 541 FVIVYFGGRREBSNNMAMVSTRLAKHIGYIELLKLMEVNPPELPQOTTKALPVRLF 600  
 Db 546 FVIVYFGGRREBSNNMAMVSTRLAKHIGYIELLKLMEVNPPELPQOTTKALPVRLF 605  
 QY 601 TDYNLSSVGETSLAEMIALTSDACEREFGLATRLPRVFKTEPTOGKKMKKTCCLPS 660  
 Db 606 TDYNLSSVGETSLAEMIALTSDACEREFGLATRLPRVFKTEPTOGKKMKKTCCLPS 665  
 QY 661 FVIFLFIIGCITIGITLLAIFRVDKHLTVNAVLISASVGLAVLNCRTMWOYDLSL 720  
 Db 666 FVIFLFIIGCITIGITLLAIFRVDKHLTVNAVLISASVGLAVLNCRTMWOYDLSL 725  
 QY 721 NSORRLNNAASKLKLKSEGFMKVLCCEVELMAMAKTIDSFONOTRLVYIIDGLDAC 780  
 Db 726 NSORRLNNAASKLKLKSEGFMKVLCCEVELMAMAKTIDSFONOTRLVYIIDGLDAC 785  
 QY 781 EODKVLQMLDTRVLESGKPFIAFASDPHIIKAINONLSVLDNSINIGHDYMRNVH 840  
 Db 786 EODKVLQMLDTRVLESGKPFIAFASDPHIIKAINONLSVLDNSINIGHDYMRNVH 845  
 QY 841 LPVFLNSGLSNARKFLVTSATNGDVPSCDTTGIOEDADRVYSONSLGEMTKLSKTA 900  
 Db 846 LPVFLNSGLSNARKFLVTSATNGDVPSCDTTGIOEDADRVYSONSLGEMTKLSKTA 905  
 QY 901 RUDYRRROMORTITROMSPDLTKLVTEDWFSIDISPTOMRLINIVSVTRRLRANOIS 960  
 Db 906 RUDYRRROMORTITROMSPDLTKLVTEDWFSIDISPTOMRLINIVSVTRRLRANOIS 965  
 QY 961 FNMDBLASMINLTEOMPRTSMLILYETEGLIPDOMLKTIERISKNIPTTADVPPL 1020  
 Db 966 FNMDBLASMINLTEOMPRTSMLILYETEGLIPDOMLKTIERISKNIPTTADVPPL 1025  
 QY 1021 EIDGDIRNEFVFLSRPVLVADVKYFLPCTVNLDPKLEIADVRAARQOISIGSLAY 1080  
 Db 1026 EIDGDIRNEFVFLSRPVLVADVKYFLPCTVNLDPKLEIADVRAARQOISIGSLAY 1085  
 QY 1081 PPLPLHEGPPRAPSGISQPSVCSSTSNFAGAGVYSPQSHSYSGMCPQHPFN-- 1138  
 Db 1086 PPLPLHEGPPRAPSGISQPSVCSSTSNFAGAGVYSPQSHSYSGMCPQHPFN-- 1145

QY 1139 -----RGSGP 1143  
 Db 1146 FFAPYLYPRYPGSGQHLISRPVATSLPRDONGLEVIKEDAAEGLSSPTDSRGSGP 1205  
 QY 1144 APGPVYLLNSLNDVAVCEKIOJEGDOSMLPOYCTTIKANKNGVLAOCNIDELKXEM 1203  
 Db 1206 APGPVYLLNSLNDVAVCEKIOJEGDOSMLPOYCTTIKANKNGVLAOCNIDELKXEM 1265  
 QY 1204 NMNFGDWHLFRSTYLENRNASHVVEDPREFLESSESGAPAPRARRASHNELPTELS 1263  
 Db 1266 NMNFGDWHLFRSTYLENRNASHVVEDPREFLESSESGAPAPRARRASHNELPTELS 1325  
 QY 1264 SQPTTYLNFSPLELNTIGLDEGAPRHSNLSWOSQOTRTPLSLNSQDSSIELSKITDY 1323  
 Db 1326 SQPTTYLNFSPLELNTIGLDEGAPRHSNLSWOSQOTRTPLSLNSQDSSIELSKITDY 1385  
 QY 1324 QAEYRDARYEYIAWMSQLEGPGSTTISGRSPHSTYYMGQSSGGSISNLEOEKGRS 1383  
 Db 1386 QAEYRDARYEYIAWMSQLEGPGSTTISGRSPHSTYYMGQSSGGSISNLEOEKGRS 1445  
 QY 1384 EPRPDGRKSFLLKRGVDIDYSSSGVSTNDASPLDITEDEKSDSGSKLLPGKKSER 1443  
 Db 1446 EPRPDGRKSFLLKRGVDIDYSSSGVSTNDASPLDITEDEKSDSGSKLLPGKKSER 1505  
 QY 1444 SSLFOTDLKLGSGLRQKLPSPDEDESGTEESONTPLDDKDKRAEGVERVPKSPHS 1503  
 Db 1506 SSLFOTDLKLGSGLRQKLPSPDEDESGTEESONTPLDDKDKRAEGVERVPKSPHS 1565  
 QY 1504 AEPTRETRAKAYELSDALDKKSDSGSVSSSSPNHSLHNEVADDSOLEKANLELBD 1563  
 Db 1566 AEPTRETRAKAYELSDALDKKSDSGSVSSSSPNHSLHNEVADDSOLEKANLELBD 1625  
 QY 1564 DSHSGKRGIPHSISGLDPLIARMSICSEDKKSPESCILASSPEENMPACOKAYLNT 1623  
 Db 1626 DSHSGKRGIPHSISGLDPLIARMSICSEDKKSPESCILASSPEENMPACOKAYLNT 1685  
 QY 1624 PSTYTLNNNAPANRANONDEMEGIRETSOVLTRSSSPNPTIONENLKSMTKRSOR 1683  
 Db 1686 PSTYTLNNNAPANRANONDEMEGIRETSOVLTRSSSPNPTIONENLKSMTKRSOR 1745  
 QY 1684 SSTYRLSKDPELHAAASSESTGFGERESIL 1715  
 Db 1746 SSTYRLSKDPELHAAASSESTGFGERESIL 1777  
 RESULT 2  
 QY 09EOG6 PRELIMINARY; PRT; 1762 AA.  
 AC 09EOG6;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)  
 DE KIDINS220.  
 GN KIDINS220.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20568256; Pubmed=10998417;  
 RA Iglesias T., Cabrera-Poch N., Mitchell M.P., Naven T.J., Rozenfurt E.,  
 RA Schiavo G.;  
 RT "Identification and cloning of Kidins220, a novel neuronal substrate  
 of protein kinase D.";  
 RL J. Biol. Chem. 275:40048-40056(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Schiavo G.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.



RA Mitchell M., Schiavo G.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF239045; AAG35185.2;  
 DR HSSP: P80144; 2MYO.  
 DR InterPro: IPR002110; ANK.  
 DR Pfam: PF00023; ank.12.  
 DR PRINTS: PRO1415; ANKYRIN.  
 DR SMART: SM00248; ANK.11.  
 DR PROSITE: PS00088; ANK\_REPEAT.10.  
 DR ANK repeat.  
 KW ANK repeat.  
 SO SEQUENCE 1762 AA; 195715 MW; 0CB2689A571F8A4 CRC64;

Query Match 94.1%; Score 8362.5; DB 11; Length 1762;  
 Best Local Similarity 91.6%; Pred. No. 0;  
 Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;

QY 1 MSVLSGSVINYEEENIPALKALECKKVDYDERNEGQTPMIAAEGMLEYKELKN 60  
 DB 1 MSVLSGSVINYEEENIPALKALECKKVDYDERNEGQTPMIAAEGMLEYKELKN 60  
 QY 61 GANCNLEDDLNMTALLASKEGHVHYEELKCGVNLNHRDMGWTALMACYKGRDYY 120  
 DB 61 GANCNLEDDLNMTALLASKEGHVHYEELKCGVNLNHRDMGWTALMACYKGRDYY 120  
 QY 121 ELLISHGANDSVTGLQYVYPIIAGRGHADIYHLLONGARVNSDKYGTTPLYWAAR 180  
 DB 121 ELLISHGANDSVTGLQYVYPIIAGRGHADIYHLLONGARVNSDKYGTTPLYWAAR 179  
 QY 181 KGHLECKHLLANGADVDOGANSMTALYAAGVGGTYQSKEILKRPNNLTDKQNTA 240  
 DB 181 KGHLECKHLLANGADVDOGANSMTALYAAGVGGTYQSKEILKRPNNLTDKQNTA 239  
 QY 241 LMAISKEGHTIEIYODLDAGTYVNIIPDRSGTYLIGAVRGHVEIYALLQKXADIDIRG 300  
 DB 241 LMAISKEGHTIEIYODLDAGTYVNIIPDRSGTYLIGAVRGHVEIYALLQKXADIDIRG 299  
 QY 301 QDNKFTALYMAVEKGNATVBDILQCNPDTEICTKDETPLIKATKMNIEVEVELLDKA 360  
 DB 301 QDNKFTALYMAVEKGNATVBDILQCNPDTEICTKDETPLIKATKMNIEVEVELLDKA 359  
 QY 361 KVSAYVDKKDPTPLHAIARGSRKLAELLNPNKDGRLLYPNKAGETPYNIDCSHOKSL 420  
 DB 361 KVSAYVDKKDPTPLHAIARGSRKLAELLNPNKDGRLLYPNKAGETPYNIDCSHOKSL 419  
 QY 421 TQIFGARHLSPTEFDGMLGYDLYSSALADILSEPTMQPICYGLYANQSGSKFLLKTL 480  
 DB 421 TQIFGARHLSPTEFDGMLGYDLYSSALADILSEPTMQPICYGLYANQSGSKFLLKTL 479  
 QY 481 EDEKRTFAGQOIEPLFOFSWLYFVLTLLCGGLGLLFAFTVHPNLGIAVSLFLALLYIF 540  
 DB 481 EDEKRTFAGQOIEPLFOFSWLYFVLTLLCGGLGLLFAFTVHPNLGIAVSLFLALLYIF 539  
 QY 541 FIYIFEGGRREGESWMAWVLSRRLARHIGYELLLKIMFVNPELPEQTKALPREFL 600  
 DB 541 FIYIFEGGRREGESWMAWVLSRRLARHIGYELLLKIMFVNPELPEQTKALPREFL 599  
 QY 601 TDVNRLLSSVGGETSLAEMIALTSDACEREGFLATRLFRVFKTEDIQGKKWKTKCLPS 660  
 DB 601 TDVNRLLSSVGGETSLAEMIALTSDACEREGFLATRLFRVFKTEDIQGKKWKTKCLPS 659  
 QY 661 FVIFLFIIGCTISGILLALFRVDPKHLVNAVILISIAVVGIAFVLCNRTMVOYDLSL 720  
 DB 661 FVIFLFIIGCTISGILLALFRVDPKHLVNAVILISIAVVGIAFVLCNRTMVOYDLSL 719  
 QY 721 NSQKRRLHNAASKHKLKSGSEPMVKCEVEILAMAKATIDSFQNTRLVYIIDGDAC 780  
 DB 721 NSQKRRLHNAASKHKLKSGSEPMVKCEVEILAMAKATIDSFQNTRLVYIIDGDAC 779  
 QY 781 EODKVLQMDTVRYLFEKGFPIAFASDPHIIKAIQONNSVLRDSNINSHGYMRIVH 840  
 DB 781 EODKVLQMDTVRYLFEKGFPIAFASDPHIIKAIQONNSVLRDSNINSHGYMRIVH 839  
 QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPCLDTGIQEDADRVSQNSLGEMTKLSKSTALN 900

DB 840 LPVFLNSRGLSNARKFLVTSATNGDVPCLDTGIQEDADRVSQNSLGEMTKLSKSTALN 899  
 QY 901 RBDYRRRQMTIRROMSFDTLKLINEMDMSDISPOTMRRLNIVSVTGLLRANQIS 960  
 DB 901 RBDYRRRQMTIRROMSFDTLKLINEMDMSDISPOTMRRLNIVSVTGLLRANQIS 959  
 QY 961 FVMDRLASWINTEQMPYTSWILYLETETEGIPQOMTLKTYERISKNIPPTKQVEPL 1020  
 DB 961 FVMDRLASWINTEQMPYTSWILYLETETEGIPQOMTLKTYERISKNIPPTKQVEPL 1019  
 QY 1021 EIDGDIRNEVEFLSRTPLVARDYKVFELPCTVNLNPKLRETIADVAREQISIGLAY 1080  
 DB 1021 EIDGDIRNEVEFLSRTPLVARDYKVFELPCTVNLNPKLRETIADVAREQISIGLAY 1079  
 QY 1081 PPLPLEGPPRPPSGSQPASCSSASFENGGPPGVSPQPSRYSGLSGHPFYRNP 1139  
 DB 1081 PPLPLEGPPRPPSGSQPASCSSASFENGGPPGVSPQPSRYSGLSGHPFYRNP 1139  
 QY 1140 -----GSG-----PAGPVVL 1151  
 DB 1140 FFAPIYTPRYRPGSQHILSRSSVYTSLPDRQNNGLPDDSGFNKQRAAVPATGSSILL 1199  
 QY 1152 NSLNDVACEKIQIEGLDQSMLEPOYCTTIKKNANINGRYLAOCNIDELKKEKNNFGDM 1211  
 DB 1152 NSLNDVACEKIQIEGLDQSMLEPOYCTTIKKNANINGRYLAOCNIDELKKEKNNFGDM 1209  
 QY 1200 SSWTVDVYCEKRLQIEGLDQSMMPQYCTTIKKNANINGRYLAOCNIDELKKEKNNFGDM 1259  
 DB 1200 SSWTVDVYCEKRLQIEGLDQSMMPQYCTTIKKNANINGRYLAOCNIDELKKEKNNFGDM 1259  
 QY 1212 LFRSTYLBKRNASHVPPEDPRFLSSSSGAPRHPHAPARAHNLPHTELSSQTPYTLN 1271  
 DB 1212 LFRSTYLBKRNASHVPPEDPRFLSSSSGAPRHPHAPARAHNLPHTELSSQTPYTLN 1271  
 QY 1260 LFRSMTLEKRSVSEQVPPEDPRFLNENSSAPVPHESARASHTLPLELSSQTPYTLN 1319  
 DB 1260 LFRSMTLEKRSVSEQVPPEDPRFLNENSSAPVPHESARASHTLPLELSSQTPYTLN 1319  
 QY 1272 FSPFEELNTGLDEGAPRHSNLSMSQOTRTPPLSLNSODSSIEISKLTDKQVAREYAY 1331  
 DB 1272 FSPFEELNTGLDEGAPRHSNLSMSQOTRTPPLSLNSODSSIEISKLTDKQVAREYAY 1331  
 QY 1320 FSPFEELNTGLDEGAPRHSNLSMSQOTRTPPLSLNSODSSIEISKLTDKQVAREYAY 1379  
 DB 1320 FSPFEELNTGLDEGAPRHSNLSMSQOTRTPPLSLNSODSSIEISKLTDKQVAREYAY 1379  
 QY 1332 REVIAQMSQLEGGPGSTTISGRSSPHSTYMGQSSSGSISNLEDEKQKSEPRPDDGR 1391  
 DB 1332 REVIAQMSQLEGGPGSTTISGRSSPHSTYMGQSSSGSISNLEDEKQKSEPRPDDGR 1391  
 QY 1380 REVIAQMSQLEGGPGSTTISGRSSPHSTYMGQSSSGSISNLEDEKQKSEPRPDDGR 1439  
 DB 1380 REVIAQMSQLEGGPGSTTISGRSSPHSTYMGQSSSGSISNLEDEKQKSEPRPDDGR 1439  
 QY 1392 KSTFLMKRGVIVYSSSGVSTNAPLDPTEDEKSDSGSKLLEGGKSSSEPSLFQIDL 1451  
 DB 1392 KSTFLMKRGVIVYSSSGVSTNAPLDPTEDEKSDSGSKLLEGGKSSSEPSLFQIDL 1449  
 QY 1440 KSTFLMKRGVIVYSSSGVSTNAPLDPTEDEKSDSGSKLLEGGKSSSEPSLFQIDL 1499  
 DB 1440 KSTFLMKRGVIVYSSSGVSTNAPLDPTEDEKSDSGSKLLEGGKSSSEPSLFQIDL 1499  
 QY 1452 KLGSGLRQKLPSPDEDESGTESDNTPLKDKDKRKAAGKVERYPKSPHSEAPETRFPI 1511  
 DB 1452 KLGSGLRQKLPSPDEDESGTESDNTPLKDKDKRKAAGKVERYPKSPHSEAPETRFPI 1511  
 QY 1500 KLGSGLRQKLPSPDEDESGTESDNTPLKDKDKRKAAGKVERYPKSPHSEAPETRFPI 1559  
 DB 1500 KLGSGLRQKLPSPDEDESGTESDNTPLKDKDKRKAAGKVERYPKSPHSEAPETRFPI 1559  
 QY 1512 KAKETYSALLDKSDSGSVSSSESSPNHSLHNEVADDSQLEKANLIELEDDSHSGRG 1571  
 DB 1512 KAKETYSALLDKSDSGSVSSSESSPNHSLHNEVADDSQLEKANLIELEDDSHSGRG 1571  
 QY 1560 KAKETYSALLDKSDSGSVSSSESSPNHSLHNEVADDSQLEKANLIELEDDSHSGRG 1619  
 DB 1560 KAKETYSALLDKSDSGSVSSSESSPNHSLHNEVADDSQLEKANLIELEDDSHSGRG 1619  
 QY 1572 IPHSLSGLODPIIARSTSEDKSSPESCLASSPEEMWPCOKAYNLNRPSTVTLLN 1631  
 DB 1572 IPHSLSGLODPIIARSTSEDKSSPESCLASSPEEMWPCOKAYNLNRPSTVTLLN 1631  
 QY 1620 MPHSLSGLODPIIARSTSEDKSSPESCLASSPEEMWPCOKAYNLNRPSTVTLLN 1679  
 DB 1620 MPHSLSGLODPIIARSTSEDKSSPESCLASSPEEMWPCOKAYNLNRPSTVTLLN 1679  
 QY 1632 NSAPANRANQNDDEMIGIRETSQVILRPSSSPNTTIOENKLSKTHKRSQSSYTRLSK 1691  
 DB 1632 NSAPANRANQNDDEMIGIRETSQVILRPSSSPNTTIOENKLSKTHKRSQSSYTRLSK 1691  
 QY 1680 NNAPTNRANQNDDEIGIRETSQVILRPSSSPNTTIOENKLSKTHKRSQSSYTRLSK 1739  
 DB 1680 NNAPTNRANQNDDEIGIRETSQVILRPSSSPNTTIOENKLSKTHKRSQSSYTRLSK 1739  
 QY 1692 DPPELHAAASSESTGGEERESIL 1715  
 DB 1692 DPPELHAAASSESTGGEERESIL 1715  
 QY 1740 DASELH-AAASSESTGGEERESIL 1762  
 DB 1740 DASELH-AAASSESTGGEERESIL 1762

RESULT 3  
 Q9ERD4 PRELIMINARY; PRT: 1715 AA.  
 ID Q9ERD4;  
 AC Q9ERD4;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Ankyrin repeat-rich membrane-spanning protein.  
 GN ARMS.

OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=20585245; PubMed=1150334;  
 RA Kong H., Boulter J., Weber J.L., Lai C., Chao M.V.;  
 RT "An evolutionarily conserved transmembrane protein that is a novel  
 RT downstream target of neurotrophin and ephrin receptors."  
 RL J. Neurosci. 21:176-185(2001).  
 DR EMBL: AF313464; AAG34167.1; -.  
 DR HSSP: P80144; 2MYO.  
 DR InterPro: IPR002110; ANK.  
 DR Pfam: PF00023; ank; 12.  
 DR PRINTS: PRO1415; ANKTRIN.  
 DR SMART: SM00248; ANK; 11.  
 DR PROSITE: PS50088; ANK\_REPEAT; 10.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 DR ANK repeat; Repeat.  
 DR SEQUENCE 1715 AA; 190532 MW; 505593C4A19A6CDC CRC64;

Query Match 93.0%; Score 8266; DB 11; Length 1715;  
 Best Local Similarity 92.7%; Pred. No. 0;  
 Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

QY 1 MSVLISQSVINVEENIPALAKLEKCKDVERNEGOTPLMTAEGNLEIKELKN 60  
 DB 1 MSVLISQSVINVEENIPALAKLEKCKDVERNEGOTPLMTAEGNLEIKELKN 60  
 QY 61 GANCMLELDLDMNTALISASKGHHVIEELKCGVNEHRDMGNTALMAACRGRTDVV 120  
 DB 61 GANCMLELDLDMNTALISASKGHHVIEELKCGVNEHRDMGNTALMAACRGRTDVV 120  
 QY 121 ELLSHGANPSTYTGQYVYPRITMAAGRHADYHLLQNGAKYNSCKYTTPLVMAAR 180  
 DB 121 ELLSHGANPSTYTGQYVYPRITMAAGRHADYHLLQNGAKYNSCKYTTPLVMAAR 180  
 QY 121 ELLSHGANPSTYTGQYVYPRITMAAGRHADYHLLQNGAKYNSCKYTTPLVMAAR 180  
 DB 121 ELLSHGANPSTYTGQYVYPRITMAAGRHADYHLLQNGAKYNSCKYTTPLVMAAR 180  
 QY 181 KGHLECVKHLAMGADVQEGANSMTALIVAGGYTSVKEILKRNPNVLTDRKDGNTA 240  
 DB 181 KGHLECVKHLAMGADVQEGANSMTALIVAGGYTSVKEILKRNPNVLTDRKDGNTA 240  
 QY 181 KGHLECVKHLAMGADVQEGANSMTALIVAGGYTSVKEILKRNPNVLTDRKDGNTA 240  
 DB 181 KGHLECVKHLAMGADVQEGANSMTALIVAGGYTSVKEILKRNPNVLTDRKDGNTA 240  
 QY 241 LMIAKSGHTEIVODLDAGTYVNIIDRSQDPTVLIGAVRGHVEIVRALLOKADIDIRG 300  
 DB 241 LMIAKSGHTEIVODLDAGTYVNIIDRSQDPTVLIGAVRGHVEIVRALLOKADIDIRG 300  
 QY 241 LMIAKSGHTEIVODLDAGTYVNIIDRSQDPTVLIGAVRGHVEIVRALLOKADIDIRG 300  
 DB 241 LMIAKSGHTEIVODLDAGTYVNIIDRSQDPTVLIGAVRGHVEIVRALLOKADIDIRG 300  
 QY 301 QDNKATLWAVEKGNATVVRDILQCNPTETCTKDGETPLKATKRNIEVVELLDKGA 360  
 DB 301 QDNKATLWAVEKGNATVVRDILQCNPTETCTKDGETPLKATKRNIEVVELLDKGA 360  
 QY 301 QDNKATLWAVEKGNATVVRDILQCNPTETCTKDGETPLKATKRNIEVVELLDKGA 360  
 DB 301 QDNKATLWAVEKGNATVVRDILQCNPTETCTKDGETPLKATKRNIEVVELLDKGA 360  
 QY 361 KVSANDKKGDTPLTAIRGRSRKLAELLRNPKDRLYRNKAGETPTNIDCSHOKSIL 420  
 DB 361 KVSANDKKGDTPLTAIRGRSRKLAELLRNPKDRLYRNKAGETPTNIDCSHOKSIL 420  
 QY 361 KVSANDKKGDTPLTAIRGRSRKLAELLRNPKDRLYRNKAGETPTNIDCSHOKSIL 420  
 DB 361 KVSANDKKGDTPLTAIRGRSRKLAELLRNPKDRLYRNKAGETPTNIDCSHOKSIL 420  
 QY 421 TQIFGARHLSPETDGMGLDYLSALADILSEPTMOPICVGLYAQMGSGKFLKL 480  
 DB 421 TQIFGARHLSPETDGMGLDYLSALADILSEPTMOPICVGLYAQMGSGKFLKL 480  
 QY 421 TQIFGARHLSPETDGMGLDYLSALADILSEPTMOPICVGLYAQMGSGKFLKL 480  
 DB 421 TQIFGARHLSPETDGMGLDYLSALADILSEPTMOPICVGLYAQMGSGKFLKL 480  
 QY 481 EDEMTKGAQOIEPLFQFSWLIIVLTLLCGGLGTFYHPNUGIAVSLFALALYIF 540  
 DB 481 EDEMTKGAQOIEPLFQFSWLIIVLTLLCGGLGTFYHPNUGIAVSLFALALYIF 540  
 QY 481 EDEMTKGAQOIEPLFQFSWLIIVLTLLCGGLGTFYHPNUGIAVSLFALALYIF 540  
 DB 481 EDEMTKGAQOIEPLFQFSWLIIVLTLLCGGLGTFYHPNUGIAVSLFALALYIF 540  
 QY 541 FIVIFGGRGEGESNMAMVLSRLRAHIGYLELLKLMFNPPELPOTTKALPVRF 600  
 DB 541 FIVIFGGRGEGESNMAMVLSRLRAHIGYLELLKLMFNPPELPOTTKALPVRF 600  
 QY 541 FIVIFGGRGEGESNMAMVLSRLRAHIGYLELLKLMFNPPELPOTTKALPVRF 600  
 DB 541 FIVIFGGRGEGESNMAMVLSRLRAHIGYLELLKLMFNPPELPOTTKALPVRF 600  
 QY 601 TDYNLSSVGETSLAEMIATLSACEREFGLATRLRVKTEDTQKKKKWKTCCLPS 660  
 DB 601 TDYNLSSVGETSLAEMIATLSACEREFGLATRLRVKTEDTQKKKKWKTCCLPS 660  
 QY 601 TDYNLSSVGETSLAEMIATLSACEREFGLATRLRVKTEDTQKKKKWKTCCLPS 660  
 DB 601 TDYNLSSVGETSLAEMIATLSACEREFGLATRLRVKTEDTQKKKKWKTCCLPS 660  
 QY 661 FVIFLITGCIISGTTLLAIFRVPKHLTVNAVLISTASVGLAFVLCNCRWQVYDLSL 720  
 DB 661 FVIFLITGCIISGTTLLAIFRVPKHLTVNAVLISTASVGLAFVLCNCRWQVYDLSL 720

DB 661 FVIFLITGCIISGTTLLAIFRVPKHLTVNAVLISTASVGLAFVLCNCRWQVYDLSL 720  
 QY 721 NSQKRLHNAASLKLKSEGMKVKLCEVELMARAKITIDSTONQOTLRYIIDGLAC 780  
 DB 721 NSQKRLHNAASLKLKSEGMKVKLCEVELMARAKITIDSTONQOTLRYIIDGLAC 780  
 QY 721 NSQKRLHNAASLKLKSEGMKVKLCEVELMARAKITIDSTONQOTLRYIIDGLAC 780  
 DB 721 NSQKRLHNAASLKLKSEGMKVKLCEVELMARAKITIDSTONQOTLRYIIDGLAC 780  
 QY 781 EODKVLMDLTVAVLSKGFIFAFSDPHIIKALNOUNSVLRNSNGHDYMRNYH 840  
 DB 781 EODKVLMDLTVAVLSKGFIFAFSDPHIIKALNOUNSVLRNSNGHDYMRNYH 840  
 QY 781 EODKVLMDLTVAVLSKGFIFAFSDPHIIKALNOUNSVLRNSNGHDYMRNYH 840  
 DB 781 EODKVLMDLTVAVLSKGFIFAFSDPHIIKALNOUNSVLRNSNGHDYMRNYH 840  
 QY 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTGTIOEDADRVSONSLGEMTKSGTALN 900  
 DB 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTGTIOEDADRVSONSLGEMTKSGTALN 900  
 QY 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTGTIOEDADRVSONSLGEMTKSGTALN 900  
 DB 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTGTIOEDADRVSONSLGEMTKSGTALN 900  
 QY 901 RRDYRRRQKRTITQMSFDLKLVTEDWEDSISFQPMRLNIVSVTGLLRANOIS 960  
 DB 901 RRDYRRRQKRTITQMSFDLKLVTEDWEDSISFQPMRLNIVSVTGLLRANOIS 960  
 QY 901 RRDYRRRQKRTITQMSFDLKLVTEDWEDSISFQPMRLNIVSVTGLLRANOIS 960  
 DB 901 RRDYRRRQKRTITQMSFDLKLVTEDWEDSISFQPMRLNIVSVTGLLRANOIS 960  
 QY 961 FNMDRLASWINTEQMPYRTSMILVLEETEGIPDQMTLKTIERISKNIPTTKDVEPL 1020  
 DB 961 FNMDRLASWINTEQMPYRTSMILVLEETEGIPDQMTLKTIERISKNIPTTKDVEPL 1020  
 QY 961 FNMDRLASWINTEQMPYRTSMILVLEETEGIPDQMTLKTIERISKNIPTTKDVEPL 1020  
 DB 961 FNMDRLASWINTEQMPYRTSMILVLEETEGIPDQMTLKTIERISKNIPTTKDVEPL 1020  
 QY 1021 EIDGDLNFEVLSSTPVLVARDVKPLPCTVNLDPKLEIADYRAAREQISIGCLAY 1080  
 DB 1021 EIDGDLNFEVLSSTPVLVARDVKPLPCTVNLDPKLEIADYRAAREQISIGCLAY 1080  
 QY 1021 EIDGDLNFEVLSSTPVLVARDVKPLPCTVNLDPKLEIADYRAAREQISIGCLAY 1080  
 DB 1021 EIDGDLNFEVLSSTPVLVARDVKPLPCTVNLDPKLEIADYRAAREQISIGCLAY 1080  
 QY 1081 PPLPIHGRPPRPSGVSOPSPVCSSTSPNPPFAGVSPQPHSSYSGLSGQHPHYNNA 1140  
 DB 1081 PPLPIHGRPPRPSGVSOPSPVCSSTSPNPPFAGVSPQPHSSYSGLSGQHPHYNNA 1140  
 QY 1081 PPLPIHGRPPRPSGVSOPSPVCSSTSPNPPFAGVSPQPHSSYSGLSGQHPHYNNA 1140  
 DB 1081 PPLPIHGRPPRPSGVSOPSPVCSSTSPNPPFAGVSPQPHSSYSGLSGQHPHYNNA 1140  
 QY 1141 SGAPGPPVLLNSLNDVACEKLOJEGIDQSMLOPYCTTIRKANINGVLAQCNIDELK 1200  
 DB 1141 SGAPGPPVLLNSLNDVACEKLOJEGIDQSMLOPYCTTIRKANINGVLAQCNIDELK 1200  
 QY 1141 SGAPGPPVLLNSLNDVACEKLOJEGIDQSMLOPYCTTIRKANINGVLAQCNIDELK 1200  
 DB 1141 SGAPGPPVLLNSLNDVACEKLOJEGIDQSMLOPYCTTIRKANINGVLAQCNIDELK 1200  
 QY 1201 KEMANNNGDHLFRSYVLEKRNASHVVEDPREFLESSESGAPRPHGPARRASHNELPRT 1260  
 DB 1201 KEMANNNGDHLFRSYVLEKRNASHVVEDPREFLESSESGAPRPHGPARRASHNELPRT 1260  
 QY 1201 KEMANNNGDHLFRSYVLEKRNASHVVEDPREFLESSESGAPRPHGPARRASHNELPRT 1260  
 DB 1201 KEMANNNGDHLFRSYVLEKRNASHVVEDPREFLESSESGAPRPHGPARRASHNELPRT 1260  
 QY 1261 ELSQTPYTLNFEFEELNTGLDEGAPRHSNLSMOSQTRTPSLSLNSQDSIEISKLT 1320  
 DB 1261 ELSQTPYTLNFEFEELNTGLDEGAPRHSNLSMOSQTRTPSLSLNSQDSIEISKLT 1320  
 QY 1261 ELSQTPYTLNFEFEELNTGLDEGAPRHSNLSMOSQTRTPSLSLNSQDSIEISKLT 1320  
 DB 1261 ELSQTPYTLNFEFEELNTGLDEGAPRHSNLSMOSQTRTPSLSLNSQDSIEISKLT 1320  
 QY 1321 DKVQAEYRDYREYIAQMSOLEGPGSTTISGRSSPHSTYYMGQSSSGSISHSLEOEKG 1380  
 DB 1321 DKVQAEYRDYREYIAQMSOLEGPGSTTISGRSSPHSTYYMGQSSSGSISHSLEOEKG 1380  
 QY 1321 DKVQAEYRDYREYIAQMSOLEGPGSTTISGRSSPHSTYYMGQSSSGSISHSLEOEKG 1380  
 DB 1321 DKVQAEYRDYREYIAQMSOLEGPGSTTISGRSSPHSTYYMGQSSSGSISHSLEOEKG 1380  
 QY 1381 KDSEPKPDQKRSKFLKRGVYIDYSSGYSTNDASPLDITTEDEKSDQSGSKILPGKKS 1440  
 DB 1381 KDSEPKPDQKRSKFLKRGVYIDYSSGYSTNDASPLDITTEDEKSDQSGSKILPGKKS 1440  
 QY 1381 KDSEPKPDQKRSKFLKRGVYIDYSSGYSTNDASPLDITTEDEKSDQSGSKILPGKKS 1440  
 DB 1381 KDSEPKPDQKRSKFLKRGVYIDYSSGYSTNDASPLDITTEDEKSDQSGSKILPGKKS 1440  
 QY 1441 SERSSSLFQNDLKLKSGSLRYOKLPDEDSGTEESNPT-LTKDKRAAEKVERVPKS 1499  
 DB 1441 SERSSSLFQNDLKLKSGSLRYOKLPDEDSGTEESNPT-LTKDKRAAEKVERVPKS 1499  
 QY 1441 SERSSSLFQNDLKLKSGSLRYOKLPDEDSGTEESNPT-LTKDKRAAEKVERVPKS 1499  
 DB 1441 SERSSSLFQNDLKLKSGSLRYOKLPDEDSGTEESNPT-LTKDKRAAEKVERVPKS 1499  
 QY 1500 PEHSAPDITRTAKAYLSDALLDKRSDSDGVRSESSSPNHSANEVADDSOLEKANI 1559  
 DB 1500 PEHSAPDITRTAKAYLSDALLDKRSDSDGVRSESSSPNHSANEVADDSOLEKANI 1559  
 QY 1500 PEHSAPDITRTAKAYLSDALLDKRSDSDGVRSESSSPNHSANEVADDSOLEKANI 1559  
 DB 1500 PEHSAPDITRTAKAYLSDALLDKRSDSDGVRSESSSPNHSANEVADDSOLEKANI 1559  
 QY 1501 QEHSAPDITRTAKAYLSDALLDKRSDSDGVRSESSSPNHSANEVADDSOLEKANI 1560  
 DB 1501 QEHSAPDITRTAKAYLSDALLDKRSDSDGVRSESSSPNHSANEVADDSOLEKANI 1560  
 QY 1560 ELEDSDHSGKRGIPHSLSGLDPTIARMSITCSDDKSPSECSLIASSPEENPAOCOKAVN 1619  
 DB 1560 ELEDSDHSGKRGIPHSLSGLDPTIARMSITCSDDKSPSECSLIASSPEENPAOCOKAVN 1619  
 QY 1561 ELEDSDHSGKRGIPHSLSGLDPTIARMSITCSDDKSPSECSLIASSPEENPAOCOKAVN 1620  
 DB 1561 ELEDSDHSGKRGIPHSLSGLDPTIARMSITCSDDKSPSECSLIASSPEENPAOCOKAVN 1620  
 QY 1620 LNRTPSTVTLNNNSANRANONFDMEGIRETSQVILRPSSSPNPTTQNTNLKSMTHK 1679  
 DB 1620 LNRTPSTVTLNNNSANRANONFDMEGIRETSQVILRPSSSPNPTTQNTNLKSMTHK 1679  
 QY 1621 LNRTPSTVTLNNNSANRANONFDMEGIRETSQVILRPSSSPNPTTQNTNLKSMTHK 1680  
 DB 1621 LNRTPSTVTLNNNSANRANONFDMEGIRETSQVILRPSSSPNPTTQNTNLKSMTHK 1680  
 QY 1680 RSQRSSYTRLSKDPPELAAAASESTGFEERESIL 1715  
 DB 1680 RSQRSSYTRLSKDPPELAAAASESTGFEERESIL 1715  
 QY 1681 RSQRSSYTRLSKDPPELAAAASESTGFEERESIL 1715  
 DB 1681 RSQRSSYTRLSKDPPELAAAASESTGFEERESIL 1715

RESULT 4  
 Q90F42

ID 09UFA2 PRELIMINARY; PRT; 1031 AA.  
 AC 09UFA2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical 115.3 kDa protein.  
 GN DKEP344F0621.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RL Blum H., Baerends S., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL133620; CAB63746.1;  
 DR HSSP; P80144; 2MYO.  
 DR InterPro; IPR002110; ANK.  
 DR Pfam; PF00023; ank. 12.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK. 11.  
 DR PROSITE; PS50088; ANK\_REPEAT. 10.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION. 1.  
 DR ANK repeat: Hypothetical protein; Repeat.  
 SQ SEQUENCE 1031 AA; 115311 MW; D9B66461C13A2F5 CRC64;

Query Match 58.6%; Score 5205.5; DB 4; Length 1031;  
 Best Local Similarity 98.5%; Pred. No. 0;  
 Matches 1008; Conservative 2; Mismatches 0; Indels 13; Gaps 1;

QY 1 MSYLISQSVINVEEENIPALKEKCDVDERNCGOTPLIAAEOGMEIVKELINK 60  
 DB 1 MSYLISQSVINVEEENIPALKEKCDVDERNCGOTPLIAAEOGMEIVKELINK 60  
 QY 61 GANCNLEDDNMTALISASKEGHVHVEELKCGVNLHRDMGWTALMAGYKGRDYY 120  
 DB 61 GANCNLEDDNMTALISASKEGHVHVEELKCGVNLHRDMGWTALMAGYKGRDYY 120  
 QY 121 ELLSHGANSVYGLQSVYPIIWAARGHADIVHLLONGAVNCSDKGTPIVMAAR 180  
 DB 121 ELLSHGANSVYGLQSVYPIIWAARGHADIVHLLONGAVNCSDKGTPIVMAAR 180  
 QY 181 KGHLECYKHLANGADVDGANSMTALIAVAGGYQTSKEILKRNPNVLTDKGNTA 240  
 DB 181 KGHLECYKHLANGADVDGANSMTALIAVAGGYQTSKEILKRNPNVLTDKGNTA 240  
 QY 241 LMIASKEGHEIYODLIDAGTYVNIPIRSGDTVLIGAVRGHVEIYRALLQKYADIDIRG 300  
 DB 241 LMIASKEGHEIYODLIDAGTYVNIPIRSGDTVLIGAVRGHVEIYRALLQKYADIDIRG 300  
 QY 301 QDNKTAIYMAVEKGNATVVDIIQCNNDTEICTKDGFTPLIKATKMNIEVVELLDKGA 360  
 DB 301 QDNKTAIYMAVEKGNATVVDIIQCNNDTEICTKDGFTPLIKATKMNIEVVELLDKGA 360  
 QY 361 KVSADVKKDPTPLHAIIRGRSRKLAELLNPNKDRLLYPRNKAAGETPNIDCSHOKSIL 420  
 DB 361 KVSADVKKDPTPLHAIIRGRSRKLAELLNPNKDRLLYPRNKAAGETPNIDCSHOKSIL 420  
 QY 421 TQITGAHLSPTEDGMGLGYDLSAALADILSEPTQOPICVGLYAKWGSGLKLLKL 480  
 DB 421 TQITGAHLSPTEDGMGLGYDLSAALADILSEPTQOPICVGLYAKWGSGLKLLKL 480  
 QY 481 EDEKTRAGOOIEPLFQFSMLIVFLTLGGLGLFAFTVHPNLTGAIVSLFALLIYF 540  
 DB 481 EDEKTRAGOOIEPLFQFSMLIVFLTLGGLGLFAFTVHPNLTGAIVSLFALLIYF 540  
 QY 541 FIVIFGGRREGESNMAMWVLTSLRANHIGYELLLKLMFVNPELDEQTTKALPVLEF 600  
 DB 541 FIVIFGGRREGESNMAMWVLTSLRANHIGYELLLKLMFVNPELDEQTTKALPVLEF 600  
 QY 601 TDYARLSVSGEETSLAMIAITLSDACEREGFLATRLFRVFKPDNQGKKKKKTCCLPS 660  
 DB 601 TDYARLSVSGEETSLAMIAITLSDACEREGFLATRLFRVFKPDNQGKKKKKTCCLPS 660

DB 601 TDYARLSVSGEETSLAMIAITLSDACEREGFLATRLFRVFKPDNQGKKKKKTCCLPS 660  
 QY 661 FVIFELTIGCIISGTTLAIFRVPKHLTVANVLISTASVGLAFVLCRTMGOVLSL 720  
 DB 661 FVIFELTIGCIISGTTLAIFRVPKHLTVANVLISTASVGLAFVLCRTMGOVLSL 720  
 QY 721 NSQRRLNNAASKLKLKSEGFPMVKACEVELMARMAKTIDFTONQTRLVIIIDGLDAC 780  
 DB 721 NSQRRLNNAASKLKLKSEGFPMVKACEVELMARMAKTIDFTONQTRLVIIIDGLDAC 780  
 QY 781 EODKVLQMLDVRVLFSPGPIIAPASDPHIIKAINONLSVLRDSNGHDMRYIVH 840  
 DB 781 EODKVLQMLDVRVLFSPGPIIAPASDPHIIKAINONLSVLRDSNGHDMRYIVH 840  
 QY 841 LPVFLNSGLSARKEFVTSATNGVPCSDITGIEDDADRVRSNLSGEMTKLSKATLN 900  
 DB 841 LPVFLNSGLSARKEFVTSATNGVPCSDITGIEDDADRVRSNLSGEMTKLSKATLN 900  
 QY 901 RRDYRRRQMQRTIRFMSFDLTKLVTEDWFSDISPQTRRLNIVSVTGRLLRANQIS 960  
 DB 901 RRDYRRRQMQRTIRFMSFDLTKLVTEDWFSDISPQTRRLNIVSVTGRLLRANQIS 960  
 QY 961 FWMDRASWINLTEQMPYRTSWLLIYLETGIPQOMTKIYER-----IS 1007  
 DB 961 FWMDRASWINLTEQMPYRTSWLLIYLETGIPQOMTKIYER-----IS 1007  
 QY 1008 KNI 1010  
 DB 1021 KSV 1023

RESULT 5  
 Q9H9E4 PRELIMINARY; PRT; 543 AA.

ID 09H9E4;  
 AC 09H9E4;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE CDNA FLJ12811 f1s, clone NT2RP2002475.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isonaga T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.,  
 RA "NEO human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK022873; BAB14285.1;  
 SQ SEQUENCE 543 AA; 59976 MW; F5D643D5A20C641D CRC64;

Query Match 31.4%; Score 2791; DB 4; Length 543;  
 Best Local Similarity 99.6%; Pred. No. 7; 6e-169;  
 Matches 541; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1173 MLPOCYTTIKKANINGRVLACQNDLKKEMMNFGDMILFRSTVLEKRNASHVYVPEP 1232  
 DB 1 MLPOCYTTIKKANINGRVLACQNDLKKEMMNFGDMILFRSTVLEKRNASHVYVPEP 1232  
 QY 1233 RFLSESSSGPAPGEPARRASHNELPHTELSSQPTVTLNFSPEELNTGLDGAHRHNL 1292  
 DB 1233 RFLSESSSGPAPGEPARRASHNELPHTELSSQPTVTLNFSPEELNTGLDGAHRHNL 1292  
 QY 1293 SWSQTRRRPSSLSLNSQSSSTFISKLTKRYAEFRDAIREYIAQMSQLEGPGSTTISG 1352  
 DB 1293 SWSQTRRRPSSLSLNSQSSSTFISKLTKRYAEFRDAIREYIAQMSQLEGPGSTTISG 1352  
 QY 1353 RSPSHSTYMGSSSGSISHSNLEOKGKDSPPKPDGKRSPLMRKGVYDVSSSGVSTN 1412  
 DB 1353 RSPSHSTYMGSSSGSISHSNLEOKGKDSPPKPDGKRSPLMRKGVYDVSSSGVSTN 1412  
 QY 1412 RSPSHSTYMGSSSGSISHSNLEOKGKDSPPKPDGKRSPLMRKGVYDVSSSGVSTN 1412  
 DB 1412 RSPSHSTYMGSSSGSISHSNLEOKGKDSPPKPDGKRSPLMRKGVYDVSSSGVSTN 1412

```

QY 1413 DASPLDPTDEDEKSDGSKLLPEKSSERSLFFONDRLTKGSGLRQXKLPSEDESGT 1472
DB 241 DASPLDPTDEDEKSDGSKLLPEKSSERSLFFONDRLTKGSGLRQXKLPSEDESGT 300
QY 1473 EESDNTPLDKDKRKAGKGYRVPKSPHSAEPIRTFIRAKEXLSDALDKKSSDSGV 1532
DB 301 EESDNTPLDKDKRKAGKGYRVPKSPHSAEPIRTFIRAKEXLSDALDKKSSDSGV 360
QY 1533 RSSSESPHSLHNEVADSOLEKANLLEEDSHSGKGIPIHSLSGLDPIIAMSICSE 1592
DB 361 RSSSESPHSLHNEVADSOLEKANLLEEDSHSGKGIPIHSLSGLDPIIAMSICSE 420
QY 1593 DKSSSECSLIASSPEENWPAKQAYNLRTPSTYTLNNNGAPNANONDEMEGIRET 1652
DB 421 DKSSSECSLIASSPEENWPAKQAYNLRTPSTYTLNNNGAPNANONDEMEGIRET 480
QY 1653 SOVILRPSSSPNPTTIONENLKSMTHKRSQSSYTRLSKDPPELHAAASSESTGEGERE 1712
DB 481 SOVILRPSSSPNPTTIONENLKSMTHKRSQSSYTRLSKDPPELHAAASSESTGEGERE 540
DB 1713 STL 1715
DB 541 STL 543

```

## RESULT 6

```

QY 09W211 PRELIMINARY: PRT: 1498 AA.
AC 09W211.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CG10074 protein.
GN CG10074.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Bokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Idegam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry B., Morris J., Mostreli A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```

```

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Styrbasas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2165-2195(2000).
DR EMBL; AE003453; AAF46710.1; -.
DR HSSP; P42773; 11HB.
DR FlyBase; FBgn0034615; CG10074.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000847; HTH_LysR.
DR Pfam; PF00023; ank; 11.
DR SMART; SM00248; ANK; 10.
DR PROSITE; PS50088; ANK_REPEAT; 9.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00044; HTH_LysR_FAMILY; UNKNOWN_1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1498 AA; 165031 MW; 7C2A888E1B64982C CRC64;

```

Query Match 26.8%; Score 2379.5; DB 5; Length 1498;  
 Best Local Similarity 37.6%; Pred. No. 6,1e-142;  
 Matches 559; Conservative 259; Mismatches 490; Indels 179; Gaps 39;

```

QY 1 MSYLISQSYINVEENIPALKALE-KKDVDERNEGQTPMLTAEDGNEIYKELIK 59
DB 23 MSLGRLRALLOYIDNNDISGLRALIDSRLTDDIDENATYVLAVAGGALFAFREFELA 82
QY 60 NGANCLIEDLDMWTALISAKKEGHVIVELIKCGVNEHBMGWTALMAACVKGRTDY 119
DB 83 RGDVQAEEDLDWMTALISAKSRNGHLDVYLLDHAQEVPHRMGWTSLMAAAYGHTEL 142
QY 120 VELLISGANPSVTSQISVYPIIWAAGRHADIVHLLQNGAKVNSDKTGTPELVAA 179
DB 143 VRLLDKGDAGNAGH-NYHIGALLWAAGRGYKDIYVELLYQKAGKAVNDKDTALVAC 201
QY 180 RKGHECVKHLAMGADVQEGANSMTALIVAKGQYQVKEILKRPNVLTPTDKONT 239
DB 202 RKNVEIVTTLKAGANVTAGMYSMTPLVLAAGGHTDCVSSILEKRPVNAADLKDWT 261
QY 240 ALMIASKEGTEIVDLDAGTYVNI PRSGDPTVIGAVRGVHEIVRALLQKYADIDIR 299
DB 262 ALCIASREGFODIILASLIAAGYVNIODRGATPLIHAKGRVVEALLKKNADVIO 321
QY 300 GQDNKTAIWAVEKGNATVRDILQCNPDTEICTYDGETPLIKATKRNIEVELLDKG 359
DB 322 GDRKTAIYTAVEKGTTPVKKLNTNPDLSATKDTGTPPLRAVRNRLIEIVHLLDRK 381
QY 360 AVYSAVDKGDPTPHIATIGRSKLAELLNRPKGRLLYPRNKGEPYVINDSHOKSI 419
DB 382 AKYTSARKGDTCLIHARAKSKTIVEALLNPKKSQLLYRANKAGETPYINDSLQYTI 441
QY 420 LQVIFGARHLSTPTEDGMDGYDLYSSALADISEPTQPPICVGLVQMGSGKSFLLK 479
DB 442 LQVYFGARLNTNNEDESGMLGYEYSSALADVLSPTLTPTTIVGLVAKMGSGKSFLLK 501
QY 480 LEDDEKTPAGQOIEPLPQFSMLIVLTL---LLCG--GLGLAFATYPMGLTAVLSISF 533
DB 502 LDEDEKNNRARQWAEPRIITSGILFTVCLVALLLTGTIVGLTWSAV-----VGSAAVGF 556
QY 534 LALLYIFVIFYFGGRGREGSNMAMVSTRARHIGVLELLKLMFVNPELDEQPTKA 593
DB 557 LLLATLLAAVAYCYQ--MDMQNAVSYQHGKEKMTPLRLILOVAFCHPFG-POSBSQA 613
QY 594 LPVRELFTDYNLSSVGETSLAEIATLSDACREGEFLATRLFRVFKTE--DTQKKK 651
DB 614 KVRPHFPAEANSASPT-IGDAVAHMLAALDAIESHWGLATRLRYRAFRPCKLVYDGM 672
QY 652 WKKTQCLSFVYFLFETICISITLLAIF-----RNDPKHLYNNAVILSISAVGLAFV 706
DB 673 WRRMCCIPVILFELATVVTGISTLVAVYTFADEREKEHILV--ALYVIAAVMGTLIC 730

```

QY 707 LNCRTMVOVLSLLNSORRLHNAASKLHKLKSEGFMYLKCVELMABAKTIDSEFTON 766  
 DB 731 THLHVLAKEVSVLSTIRIVLAKRAV-----RSSSEAPVLTGAVVMTDVKCLDAFTNO 786  
 QY 767 QTRLVVITDLDACEODKVLQMLDFAVRLVS--KGFPIAFASDPHIIKAIQONNSVL 824  
 DB 787 QSRVGVYDALDSCDFRITLTLNAVOTLLSPNRFVLLISDPHVIAKAAANRRLE 846  
 QY 825 RDSNNGHDYMRNIVHLVPLNSRGLSNARKFLVTS-----ATNGVPCSD--TTGIQED 877  
 DB 847 TEGGIGHGDFLRNIVHLVPLVYQNSGRKRYQRAQMTALLFKRSGGGYQTDGDPGLGHSVS 906  
 QY 878 ADDRVSQNS--LGEMFL-----GSKTALNR-----RTYRRQMQQRTITQMSR 920  
 DB 907 A-RRLSNASELISQOEKLRGPANGGKKLRLSESVASSTGSLNHRGQMPQTV-----L 960  
 QY 921 DLTKLVTEDEMFSDISPOTMRRLNTVYVGRLLRANQISFNMRLASWINTLEQMPYRT 980  
 DB 961 DLSRIYVLTDDYFSDVNRSMRLMNVYIVRLKKAQLEFESVYRLSSWINTLEQMPYRLA 1020  
 QY 981 SWLLIYEE--TEGIPDQMTLKTIERISKNIPTTDVPELLEIDGDIRNFEVFLSSRTP 1038  
 DB 1021 SMIVLHHDQPMDSNADESYSLSQSYEKLREKLAYLRBAAPLLELDREKRLDAFLQHS 1080  
 QY 1039 VLVARDYKVPVLPCTVNLDPKRLRELIAD-----VARAREQISIGGLAVPPLME 1087  
 DB 1081 DLVADIRITFLPTINIDPYLKVYKEDQOTIEDEGSLYQAKRVSNTMKRFP----- 1134  
 QY 1088 GPRAPSGYSQPPSVCSSTSFNGPFAGVVSPQPHSSY-----YSGMTG 1131  
 DB 1135 -----APTYY-----VPSQAPYPPYQFQNEYEPANELRSNLSSTTE 1171  
 QY 1132 PPHFVYRSGGPAPGPVY--LNSLVNDANCEKLOEGEDQSLMPOYCTTIRKANINGR 1189  
 DB 1172 PVPPLNPSDFSFGDILLQTKLTDLYEGVISLDRHEDM--KRALPKLAVLREMNINGR 1230  
 QY 1190 VLAQCNIDELKKEKMNFGMDHFLFRSTVLEMRNAESHVVP-----EDPRFSESSEGP 1242  
 DB 1231 VIKHCDMPDLKSVGLSGFWELFRLLITLRECE--LPRKQORQOQCGALEAPSNVP 1288  
 QY 1243 APHG-----EPARR--ASHNELPHTE-----LSSQ-----TPYTLNFSFEELNTGLDEGA 1286  
 DB 1289 MIKQVTDALMQPREPSLRKNSVSHMEKQVLEQMLIGTILQLENEAYE--DVASSERP 1347  
 QY 1287 PRHSNLSWQOTRTTPELSL-----NSQDSIEISKLTDKVQ--AY--KDAYREYI 1335  
 DB 1348 PTGEMLAAYAOQLAPLRESSESPSPDDQKQYGVKISNNNNNQYLHAERYNSVSSHSL 1407  
 QY 1336 AQMSOLEGPGSTTISGRSPHSTYYWGQSSSGSISNLEQEKGD 1382  
 DB 1408 QSLSTLVGAP-----VGHGSGSG--SHHLGNGND 1435

RESULT 7  
 Q20109 PRELIMINARY; PRT; 1398 AA.  
 AC Q20109;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE F36H1.2 protein.  
 GN F36H1.2  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditiida; Rhabditioidae;  
 CC Rhabditiidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RA SEQUENCE FROM N.A.  
 RA Keshaw J.K.  
 RL Submitted (Jan-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT Investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; 268760; CAA92996.1; -.  
 DR HSSP; 000420; IAWC.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR001412; trna-synt\_L.  
 DR Pfam; PF00023; ank; 10.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 9.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_L; UNKNOWN\_1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 7.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 KW ANK repeat; Repeat.  
 KW SEQUENCE 1398 AA; 156568 MW; D5080A9E89AB1BEC CRC64;  
 SO

Query Match 20.6%; Score 1832.5; DB 5; Length 1398;  
 Best local similarity 30.9%; Pred. No. 3.2e-107;  
 Matches 468; Conservative 280; Mismatches 483; Indels 285; Gaps 40;

QY 34 RNECGTPIIMIAEAGNLEIYELIKNGANCLWEDLN--WTALISAKEGHIVIEELL 91  
 DB 58 RNANGSELTLTVASRGNTAVAKQALDLPDA-IDETDEGWSALINAAHCGHVDIVRLI 116  
 QY 92 KCGVNLHHDGNGTALMAACCKGRDVAVELLSHGANSVYGLQYVPIIMAGRGA 151  
 DB 117 DNGASVDOPDLMGWSPDLMAVYKKNHLDVVDLVN----- 150  
 QY 152 DIVHLLHQGAVNCSDKYGTTPLYWAARKGHECVKHLAMGADVDEGANSMTALIVA 211  
 DB 151 -----AKVNCMDKFGSTALIMAAKRGHLPVQQLLNGAVDAVGMVSSATMLA 200  
 QY 212 VNGGTQSYKEILKRNPNVNLTDKQNTALMAKSEGTIEYODLIDAGTYVNIIDRSGD 271  
 DB 201 TRGNFQVVELLTLREPNVNAVADNGLFALGMAARDGVADICESLNSGAFVNOCDRGN 260  
 QY 272 TVLIGAVRGHVEIVALLQYADIDIRGQDKTALYMAVEKGNATVYDIDOCPRDEI 331  
 DB 261 WITTSVRSNGMAIYMLIDKFRADINCOSEKRTPLHLAIDSFNDIAYILLEKPNLEL 320  
 QY 332 CTKDETPLIKATKMRNIEVEVELLDKGAQVSAVDKGDPTPLHAIIRGRSLAELLRN 391  
 DB 321 KMKDGETPLBRAKACHVHLCTYIMSEFAGKLAAYVNCBGMALHLALRAASRLTQALLSN 380  
 QY 392 PKDGLLYPNKAGETPYNIDCSHOKSLITQFGARHLSPTEDGM--LGYDLYSAL 448  
 DB 381 PSDSRLLYRPNKIGQTPYSIDISNPPIPLIFG-----PIDAEKMDTAMGYDYVSVL 435  
 QY 449 ADLISEPTMOPPCVGLVYQWGSKSFLLKLEDEKKTFRAGQIEPL-FQFSLIVPLTL 507  
 DB 436 ADIVCEPSLSLPTLIGLVAKWSGSAKALAKLEAHSHSRWMLDGVLSVSFFALFALF 495  
 QY 508 LLCGGGLLFAFVHPNLGIA--VLSFLALYIF-----IVYFGGRREGESNNWAV 560  
 DB 496 LFGMFSLEFMTMIAISNSVTAYLLISWSVFLIFIFGLIYVYVYGGDK--NNYTSMD 552  
 QY 561 LSTRRLRHIGYLELLIKMFVNPELPEDDTALPVRLEFTDYNRSLSSVGETSLAEMIA 620  
 DB 553 IANFARVRSRIRLIVNITLHAP--MNSDSSAMPVSFLFADYHRLSSIGGQALAKIYA 611  
 QY 621 TUSDACEREGFATLRFVFKEDQGGKKKKKCCPLSPV-----IFLFIIGGIGSI 675  
 DB 612 TLFEAEETHHGVLPVRLFCOMKPPYGIHGLSRRCGVHVIILLIYAVFLIIMAOYFGIV 671  
 QY 676 TLIAIFRVDPKHLTVNAVLIISAVYGLAVLNCRTMVOVLSLLNSORRLHNAASKLH 735  
 DB 672 WLLS--DROPNNN--LFLAIAFLCGFMIAIYPLALIMYSMTNVPRRYNAARNAH 726  
 QY 736 KLSSEGMKYLKCEVELAMAKTIDSFQONOTRLVYIIDGACQDQAVLQMDTVRYL 795  
 DB 727 KIRFEGIMOKIQTEVLDLADMIRSDAFRSHRLVYVVDGDNCEQEBRWQVTLALIELL 786





ID	099407	PRELIMINAR:	PRT; 1956 AA.
AC	099407;		
DT	01-MAY-1997 (TREMBlrel. 03, Created)		
DT	01-MAY-1997 (TREMBlrel. 03, Last sequence update)		
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)		
DE	Ankryn.		
GN	ANK1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-97382244; PubMed-9235914;		
RA	Gallagher P.G., Tse W.T., Scarpa A.L., Lux S.E., Forget B.G.;		
RT	"Structure and organization of the human ankryn-1 gene. Basis for		
RT	complexity of pre-mRNA processing."		
RL	J. Biol. Chem. 272:19220-19228(1997).		
DR	EMBL; U50133; AAB47805.1; -		
DR	EMBL; U50092; AAB47805.1; JOINED.		
DR	EMBL; U50093; AAB47805.1; JOINED.		
DR	EMBL; U50094; AAB47805.1; JOINED.		
DR	EMBL; U50095; AAB47805.1; JOINED.		
DR	EMBL; U50096; AAB47805.1; JOINED.		
DR	EMBL; U50097; AAB47805.1; JOINED.		
DR	EMBL; U50098; AAB47805.1; JOINED.		
DR	EMBL; U50099; AAB47805.1; JOINED.		
DR	EMBL; U50100; AAB47805.1; JOINED.		
DR	EMBL; U50101; AAB47805.1; JOINED.		
DR	EMBL; U50102; AAB47805.1; JOINED.		
DR	EMBL; U50103; AAB47805.1; JOINED.		
DR	EMBL; U50104; AAB47805.1; JOINED.		
DR	EMBL; U50105; AAB47805.1; JOINED.		
DR	EMBL; U50106; AAB47805.1; JOINED.		
DR	EMBL; U50107; AAB47805.1; JOINED.		
DR	EMBL; U50108; AAB47805.1; JOINED.		
DR	EMBL; U50109; AAB47805.1; JOINED.		
DR	EMBL; U50110; AAB47805.1; JOINED.		
DR	EMBL; U50111; AAB47805.1; JOINED.		
DR	EMBL; U50112; AAB47805.1; JOINED.		
DR	EMBL; U50113; AAB47805.1; JOINED.		
DR	EMBL; U50114; AAB47805.1; JOINED.		
DR	EMBL; U50115; AAB47805.1; JOINED.		
DR	EMBL; U50116; AAB47805.1; JOINED.		
DR	EMBL; U50117; AAB47805.1; JOINED.		
DR	EMBL; U50118; AAB47805.1; JOINED.		
DR	EMBL; U50119; AAB47805.1; JOINED.		
DR	EMBL; U50120; AAB47805.1; JOINED.		
DR	EMBL; U50121; AAB47805.1; JOINED.		
DR	EMBL; U50122; AAB47805.1; JOINED.		
DR	EMBL; U50123; AAB47805.1; JOINED.		
DR	EMBL; U50124; AAB47805.1; JOINED.		
DR	EMBL; U50125; AAB47805.1; JOINED.		
DR	EMBL; U50126; AAB47805.1; JOINED.		
DR	EMBL; U50127; AAB47805.1; JOINED.		
DR	EMBL; U50128; AAB47805.1; JOINED.		
DR	EMBL; U50129; AAB47805.1; JOINED.		
DR	EMBL; U50130; AAB47805.1; JOINED.		
DR	EMBL; U50131; AAB47805.1; JOINED.		
DR	EMBL; U50132; AAB47805.1; JOINED.		
DR	HSSP; Q00430; IAWC.		
DR	InterPro: IPR002110; ANK.		
DR	InterPro: IPR000488; Death.		
DR	InterPro: IPR000906; ZUS.		
DR	Pfam: PF00023; ank; 23.		
DR	Pfam: PF00531; death; 1.		
DR	Pfam: PF00791; ZUS; 1.		
DR	PRINTS; PR01415; ANKRYIN.		
DR	SMART: SM00248; ANK; 21.		
DR	SMART: SM00005; DEATH; 1.		
DR	SMART: SM00218; ZUS; 1.		
DR	PROSITE; PSS0088; ANK_REPEAT; 20.		
DR	PROSITE; PSS0129; ANK_REPEAT_REGION; 1.		

[illegible]



```

Db 913 IPRPCAPRTITCRLVQKLTSTPPPLAEFEGGLASRIALGPTAQLSTPIVIEIPIHA 972
QY 846 -NSRGLSNARKFLVYSATNGDVPSCSTTIOEDADRRYSONSIGEMTKLGSKTALNRDPT 904
Db 973 SHGRG--DRELVVLSKENGSV-----WKHSRSGESYLDIOLN-GMEELGSLSE 1020
QY 905 YRRROMQRTITRQMSFDLTKLVLT--EDWFSDISPQMRRLNLTVSYGRLLRANQISF 961
Db 1021 LKKRRCRLIT--TDEPLFYVMSKLCQD-YIITPEG-----GSLSKVLPLVQATF 1070
QY 962 NMDRLASWINTLEQMPYRTSWLILYEETGIPDQMTK-----TIYERISK- 1008
Db 1071 PENNAVTKRYKALQ-----AQPYDELVTKLGNQATFSPITYVEPRRRKF 1116
QY 1009 -----NIPTKDVEPLEIDGDIRNEVF-----LSKRPVLYARVYK 1046
Db 1117 HRPGLRLPLPPSWNDPNRDSGGDTSLRLCSYIGTDOAQWEDITGTTKLVYANCSA 1176
QY 1047 VELPCTVMDPLREIADVRAARQISIGLAPPLPLHSGPPRAPSQSPSCST 1106
Db 1177 NF---TTNV--SARFMLSQCPRTAFAVNFATILYEL-----TAV 1211
QY 1107 SENGFPAGGVSPQHSSTYSQMTGPQHPFTNRKSGPARGPV--VLNSLVNDAVCE-- 1161
Db 1212 PYMAFV-----IFAKMNDPREGRLKCYCMTDDKYDVKTLQHE 1249
QY 1162 -----KLQIEGLDQSMUPQYCTTIKANINGRVLACINDELKEMNNFMGMHLFRS 1215
Db 1250 NFVEVARSRDIEVLEGMEL-----FAELSG-----NLVPYKAAQORSFHFQSPRE 1295
QY 1216 TVLEKRNMAESHVPPDPPLSESSGSPAPHEPARASHNE--LPHTELS----- 1263
Db 1236 NPLAM-----PKAVRDSREPGSLSFLKAKMYEDTOHILHNTWTPCKAG 1344
QY 1264 -----SQPYTINSEFEELNTIGLDEGAPR-----HSNLSWQOSTRTP-SL 1304
Db 1345 SGAEDRRRPTPLARYSILSESTPGSLSGTEQAKMVAVISEHLGSLWAELELOFSV 1404
QY 1305 SLSNLSQDSISIEISKLTQVQAEYRDAYREYIAQMSOL-----EGGPGSTTIGRSSPSTY 1360
Db 1405 EDIN-----RIRVENPRSLLEQSVALLNLVIREG-----QANMENLY 1443
QY 1361 YMGSSSGGSIHSNLEQEKGDSEPKPDGRKSFMLRKGVIDYSSSGVSTNDASPLDPT 1420
Db 1444 TALSDIGKEIVNMLESGRQSRNLKPD-----RRHTDRDYSLSPSQNGYSSL-- 1492
QY 1421 TEEDE-----KSDQSGSKL-----LPQKSSERSLSLQTDKILKSGSLRY 1460
Db 1493 --QDELLSPALGCLASPLRADQWNEVAVLDAIPLAATHEHDMLEMSDMQVVASGLTP 1550
QY 1461 QKLPSDEDE--SGTEESDNTPLLKDKRKAKEGKVERKVPKSPHSAEPIRTFKAKEYL 1517
Db 1551 SLVTAEDSLSCSKAEDSDAT-----GHEWKLLEGALSEPRGPE-----LGSLELV 1596
QY 1518 SDALLDKKDS-----DSGVASSSESPNHSIHNVEYADDQLEKANLLELEDSDHSGK 1569
Db 1597 EDDTVDSDATNGLIDLEQEGGORSSEKRLPGSKRQDQDATGAGQ--DSENEVSIVSGHORQ 1655
QY 1570 RGIFHSLSGLQDPIIARMSISEDK-----KSPSECSLIASSPRENN 1611
Db 1656 ARITHS-----PVSQVTERSQDRIDQWMDADGSIYSIYLODAAGSW 1696

RESULT 11
ID 08YTG9 PRELIMINARY; PRT: 426 AA.
AC 08YTG9:
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein A112748.
GN A112748.

```

```

OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriuguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsuno M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT *Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.*;
RL DNA Res. 8:205-213(2001).
DR EMBL, AP003590; BAB74447.1; -.
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ank.13.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK.13.
DR PROSITE: PS50088; ANK_REPEAT.11.
DR PROSITE: PS50297; ANK_REGION.1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 426 AA; 43800 MW; CF0B9B5D43ACAA03 CRC64;

Query Match 6.1%; Score 544; DB 16; Length 426;
Best Local Similarity 37.2%; Pred. No. 3.3e-26;
Matches 149; Conservative 64; Mismatches 180; Indels 8; Gaps 6;

QY 9 VIVTEENIPALKALLECKDYDERNEGQPTPIAABOGNLEIYKELIKGANCNT-E 67
Db 8 LKVAKSGDIKGLGALLAGVGVDICDRGTTALMPAAVTEIYRSILDDGAVNVAR 67
QY 68 DLDNNTALISAKREGHVHYEELKGVNLEHNDMGRTALMACKGTDTVEILLISG 127
Db 68 KRYGLRALMAASANDVYVOLLISRGAAVNATNEDSGSLAALAKGNEVAVRYLLAG 127
QY 128 ANPSVTGLQYVPIIWAAGRGHADIYHLLONGCAVNCSDKYGTPTLYMAARKHLECY 187
Db 128 ADVNITDKDDDT-ALKLAVKRGQAAYVOLLPSGADANDEDEGETILLMLADSGHDV 186
QY 188 KILLAGADVDDEGANSMTALIVAVGTYOSKEYILKKNPNVNTDKDGNALMASKE 247
Db 187 QVLLAAGVVDNEODGGTALLAAVAGCAIAKILLDAGVAVNHODDGSBSILHATVE 246
QY 248 GHTEIYODLDAGTYVINIDRSQDYLIGAVNGHVEIYRALLOKAYADIDIGQD-NKTA 306
Db 247 GYVDVQVLLNGCANTQINKLGDTPLLVLAALQGHQIVETIL-KTG-ANVHGNDVETP 304
QY 307 LYWAVERGNATVRYDILQCNPTETICTKDGETPLKATKRNIEVELLDGKAVSAD 366
Db 305 LTLAASQGHATVRYDILDGANANIPASQKTAIKATERNPVYQLLLAGAVVNTQD 364
QY 367 KKGDTPLHAIKRSKRIKAEILLRNPKDRLLYRPKAGET 407
Db 365 SVGATALLMAASGGYNKVYVILLLEGADTNL--KNRGYT 402

RESULT 12
ID 061307 PRELIMINARY; PRT: 1943 AA.
AC 061307; 061305; 061306; 061308; 061309; 061310; 008866; 008867;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ankyrin 3 (Ankyrin G) (epithelial ankyrin) (Ankyrin-3).
GN ANK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1-6).
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;

```



```

Db 666 LLSRANANVLNKNKSLTFLHLAEDRVNAEVLNAGAHVDAQTKMGYTFPLHNGCH--Y 723
OY 470 GSGK--SFLTK--KLEDEMT--FAGQO--IEPLQPSMLVLETLILCGGL 513
Db 724 GIKIVNPLLOHSAVNAKNGYALHQAQOQHIIINVLNN-- 769
OY 514 GLLFAFVHPNIGVLSLFLALIFYFIVYFGRRGESMNAMVSTRLARHIGYLE 573
Db 770 ASPNELTYNCTALNLI-- 793
OY 574 LILKMFV--NPPPLPQCTKALPVRFLFYDYNLSVSGGETSLAEMIA 620
Db 794 VDTLKVYTEEIMTTTTITEKKNVPEPMNEVID--MSDEVKAKASAPKESDGEYI- 849
OY 621 TLDACERE--FGF-LATR--LRFVFTED--T 646
Db 850 --SDGEEDATIGTDKYLGPQDLKELGDSLPAGEYVGSILARSASLSFSDSRYT 907
OY 647 QGKKKKKTCCLPSFVIFLITIGTILLLAIFRPDP-KHLT--VANVLISI 697
OY 908 LNRSSYAR--DSMMIEELLVPSREQHLTFREFSDSLRHYSMAADTLDNVNLVSSPV 963
OY 698 ASVGLAFVLCRTWQVYDLSLNSQRKLNMA--KLKL- 737
Db 964 HSGFLVSPWVDRG--GSMRGRHGMRIIPPRKCTAPRITCRLKRRKLA 1014
OY 738 --KSEG--FMVKCEVELAMAKATIDSPFN 766
Db 1015 NPPVVEGEGLASRLVEMGPAQOFLGPVIEIPHGSMRKRELLIVLSENETWKEH 1074
OY 767 Q--TRLVYIIGLACEDQKVLQMDIVR--LFSKGFIAFASDPIIITKAINON 819
Db 1075 QFDSKNEDLALINGMD--EELDSPEELGTGRICRIITK-DEPOYFA--VYSRIKOE 1126
OY 820 LNSVLRDSNINHDMYRNIVHLPELNSRGLSNARKFLY--TSANGDVPDSDTTQI-- 875
Db 1127 SNOIGPEGI--LSSTVPLVOASPEGALTRIRNGVLAOP 1166
OY 876 --EDARRVSONSLGEMTKIGSKTALNRDITY--RRROMOTITROMSFDLTKL--YT 928
Db 1167 VPETVKKI--LGNKATFSPIVTVEPRRRKFKPLT--MTIPVPPSGSGVS 1214
OY 929 EDWSDSIPQMRRLNIVSYTGRLLANQISFWMDLASVINTEMP--YR 979
Db 1215 NGYGDATP-NLRLLCSITGTSF--AQMEDITGTTPLTFIKDCVSFT 1259
OY 980 TS--WL--ILYLETEGIPDQMTLKI-YERISKNIPPTKVELELDGDIR- 1027
Db 1260 TNVSARFWLADCHQVL--ETVGLASQLYRELICVPYMAKFVYFAKTNDP--VESLRC 1313
OY 1028 --NEFVFLSRTPVLVARDVYKFLPCTVMDPLR--ELIADVRA 1068
Db 1314 FCMOTDHDVDTLEQOENFEVANSK-DIEVLEGGPIYDVCYNLAPLTKGGQGLVNFYS 1372
OY 1069 AREQISIGLAPPLPHEGPPRPSGY--SQPSVSSSIFNCPFGVYVSQPHS 1123
Db 1373 FKEN--RLPFSIKTRDTSQEP--CGRLSF-- 1397
OY 1124 SYSGMTPOHPFYNRSGPAPGVVLLNSLNDVACEKLQLEGDOSMLPQYCTTIK 1183
Db 1398 --LKEPKTKGLPQIV-- 1412
OY 1184 ANINGRYLACNIDELKEKNMNFQDMLFRSVLEKRNASHVVEDPRLSESSGPA 1243
Db 1413 CNINITLPAHKKAKEKDRROS--FASIALKRRYS--YLPEPMSPO 1454
OY 1244 PHGPPARR--ASHNELPHTELSOTPYTLNPSPELNTLIGDEGAPRHSUS-- 1293
Db 1455 --SPGRDIRMAIYADHIGLSWTELARE--LNFVDEINQIIVEN--PNSLISOS 1504
OY 1294 --W--QSOYTRTPSLNSQDSSEISLTKDVOAE--RDAYREYIAQM 1338
Db 1505 FMLLKVVTRDGNATDALTSVLTIKINRDIYTLLEGPIFDGNGISGRSFADENNVH 1564

```

```

OY 1339 SQLECGFG-----STTISGRSSPASTYTYMGQSSGGSINS--NL 1375
Db 1565 DPVDGHPSPFOVELETMPGLTYTPNPFOQDHFSDISIESPRTPSRLSDGLVPSQNI 1624
OY 1376 EQEKG-----DSS--EPKPDGKRSFLMKRGVYIDYSSGVSTNDASPLDPITEED 1425
Db 1625 EHPPGPPVYATDTSLEDSKMD-----SVTVDPDAPLD--VDESQ 1665
OY 1426 KSDQSGSKL-----LPKKSSESS--LFQTDLKIGSGLYOKLPDEDESGTE-- 1473
Db 1666 LKDLQSECAQCAVAPGIPNDGRQAEPRLPQRYKMGSEQEKKSPPDEYVEDKTK 1725
OY 1474 -----ESDNPPLKDKDKRAEKVERVP-----KSPHSAB-PITFETKA 1513
Db 1726 SLEFDIQLEVEAEEMTEDDQAMLRVORAEIAMSISLAWONETPSSGLESPPAQ-- 1781
OY 1514 KEYISDALDKKSDSGVRS-----ESSPNHSLHNEVADDSQLEKANLIEL 1561
Db 1782 -RLTGLGLDRDDSSDQARDSTSYLTGEPKITEANGHTA--EYIPRA--KAKPYFP 1835
OY 1562 EDDSHSGKGI-----PHSLSGLDPIIARMSICSEDKKSPSECS--LIASSPEMWP 1612
Db 1836 ESQNDIGKOSIKENLKPXTHGCGRTPEPV-----SPLTAYOKSLETSKLVIEDAPKCPV 1891
OY 1613 ACQKAVNMRPTVTLNNSNAPANRANONPDEMIGIRTSQVILRPSSPMTIIONEN 1672
Db 1892 VGKK--KMTRTYAD-----GKRLNLOEES--STRSEPK--QGBG 1926
OY 1673 LKSMTRK 1679
Db 1927 YKVKTKK 1933

```

```

RESULT 13
O8VC68 PRELIMINARY; PRT; 1726 AA.
ID O8VC68
AC O8VC68;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 188.2 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary Gland;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC021657; AAI21657.1; -
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; Z05.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; Z05; 1.
DR PRINTS: PR01415; ANKRYIN.
DR SMART: SM00248; ANK; 22.
DR SMART: SM00218; Z05; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW Hypothetical protein.
SQ
SEQUENCE 1726 AA; 188239 MW; 43014867500B1CA CRC64;

```

```

Query Match 5.9%; Score 522.5; DB 11; Length 1726;
Best Local Similarity 19.8%; Pred. No. 8,4e-24;
Matches 406; Conservative 246; Mismatches 612; Indels 789; Gaps 82;
OY 14 EENIPALKALKCKDQVDERNECGQTPPLMTABQGNLEIVKELIKNGANCNLELDLNMWT 73

```

```

Db      66 KEGHEVYSELQREAVNATKKGWTFALHIALSLAQAEVYKLVANGANVANNOQNGFT 125
QY      74 ALLSASKEGHHVYEELKGVNLEHRDMGWTALMACYKGTIVYELL----- 124
Db      126 PLVMAQENHLEVRFLDNGASQSLATEDGFTPLVALOQGHQVSLLENDTKGYR 185
QY      125 -----SHGANDSYGL----- 135
Db      186 LPAHLIAKBDPTKRAALLLQNTNADVESKSGFTPLHIAHGINVATILLNRAAYD 245
QY      136 ---QSYVPIIIMAGRHADIVHLLONGAKVNSDKYGTPLVMAARKG----- 182
Db      246 FTARNITPLHASKGNANMVKILLDRGAKIDAKTRDGLTPLHCGARSGEQVEMLLD 305
QY      183 -----HLECVHLLAMGADVNOEGANSATLIVAKGYTOS 219
Db      306 RSPAPILSKTKNGSLPHMATQGDHLCVOLLQHNVPVDVNTDYLTALHVAHCHGYK 365
QY      220 VKEILKRNPNVNLTDKDGNTALMASKE----- 247
Db      366 AKVLDDKXKSPNKKALNGFTPLHACKKNIRYVELLKHGASIQAVTESGLTPHYANF 425
QY      248 -GTEIYODLLDAGTYVINIPDRSGTVLIGAVRGHVEIYRALLQRYADIDIRGDNKTA 306
Db      426 MGHVNIYSQLMHGHGASPNNTNGETALHMAARSQAEVYRYLVODQAEAKAKDQTP 485
QY      307 LYMAVEKGNATWTRDILQCPDTEICTNGEPTLTKMRNIEVEVELLDKAKAYSAD 366
Db      486 LHLSARLGRADIVQOLLQOGASPNAAFTSGTPLHIAARGHEDVAFLDHCASISTTT 545
QY      367 KKGDTPLHIAIRGRSKLAELLRNPKDGLRYPKAGETPYNIDSHOKSILTOIFGA 426
Db      546 KKEFTPLHVAKYGKLEVASILLQKSAS-----PBAAG-----KSLTPLHVA 588
QY      427 RHLSPTETDGMGLYDILYSALADIISPTMPPRCVGLAQMGSKSLKLEDEMKT 486
Db      589 AH-----YNOKVAL-----LILDOGASPHAAK----- 612
QY      487 FAGOEPLRFOFMSLIVFTLLCGGLLFAFTVHPNIGIIVSLALYFFIYIYF 546
Db      613 -----NGTPLHIAAKNOMDIATSL-----LEI 636
QY      547 GG-----RRGESWNNAMVLSRIARHIGYLELLKIMFVNPELPEOTTKALPVREFT 601
Db      637 GADANAVTROG-----IASVHLAAQEGHVADVSLLSRN-----A 671
QY      602 DYNRSSVGETSLAEKIATLSDACEREGFLATRLFRVFKTE-----DTQCKKKW 652
Db      672 NVN-LSNKSGLTPL-----HIAAOEDRVNVAEVLVNOGAHVDAQTKMGY 714
QY      653 KTRCCLPSEVIFLEIIGCIISGITLLAIFRVPDKHLTVNVAVLISIAVYGLAVLCRTM 712
Db      715 -----TPPHV-----GCHGNKTI-----VNFLOHSAKV-----NAKT- 743
QY      713 MOVLDSLNSORRLHNAASKILKSEGMVKLKEVELMAMAKTIDISFTONOTRLVY 772
Db      744 -----KNGYTAHLHQAQGHHTIIT--LLONNASPNELTVNGNTALAI 785
QY      773 IIOGLACBQDKVLOMDIVRYLFSKGPFLAIFASPHIILIKAINONLSVLRDSINCH 832
Db      786 -----ARRLGIYSVDTLKVY-TEIIMTTTTITEKHKM--NVPETMEVLDMSDECE 835
QY      833 DYMRNIYHLPVFLNSKGLSNARKFLVTSATNGDVPDSTGIEODADRYSQSLGEM-- 890
Db      836 D-----ALTGDDTKVIGPODLKELGD 856
QY      891 -----TKLSKTA-LNRDRTYRRROMORTITROMSFDLTKLV-----TED 930
Db      857 DSLPABEGYVGLSAGASASLRSFSDRSYTLNRSYARDSMMIEELLVPSKEOHLFTFE 916
QY      931 WPSDISPQMRRLNLIYVYGRLLRANQISFNW--DRLASWILUWCOMPEKTSWILLYLE 988

```

---

```

Db      917 FDSD-----SLRH-----YSMAADTLDN-VNLVSS-PVHSGFLSEMV 952
QY      989 ETEGIPDQMTLTIYERISKNIPTKYDVEPLLEIDGDRNEVEFLSSFTPLVADKYVF 1048
Db      953 DARG-----GSMR-----GSR-----HHGMRI 970
QY      1049 LECTVNLDPKLEILLIADYRAAREOISIGLAVPPLPLEGEPRAPSGSOPEVSSSTSF 1108
Db      971 IP-----PRKTAPTRTICRLYKRIKLANP--PMVEGEGSLASLYEMGA----- 1014
QY      1109 NGPFAGVSPQPHSSY-----SGMTPOHPFYNRGSGPAPVYLLSLN 1155
Db      1015 GAQFLGAPVLEIPHFSGMSMGKEREILVLRSENGETWKHQFDKNEDLAE-----LLNGMD 1070
QY      1156 VDAVEKRLQOIEGLOSMU-----POYCTTIKANIINGRYALACNIDELKEMNNF 1207
Db      1071 -----EELDSPELCTKRICRIITTKDFPOYFAVNSR-----IKOESNOIG 1110
QY      1208 GDMHLFRSTVLEMRNA-----ESHVVPED--PRFL-SESSGAPAPHGPA 1249
Db      1111 PEGGLISSTVPLVQASPEEGALTKRINVGLOAQVPEPEYKILGNKATSPITYVEPR 1170
QY      1250 KRASHNELPHT-----ELSOQTP-- 1267
Db      1171 RRFKFKPITMTIPVPPSGEGVNGKGDATFNLRLCSITGTSIPAQMEDIOTTPLE 1230
QY      1268 -----YTLNFS-----PELNTLIGDEBAPRHSNLSMOSQTR--TPSLSL-- 1307
Db      1231 IKDCVSPFTNVSARFWLADCHOVLETVG-----ASOLYRELICVPIYAKRVV 1278
QY      1308 -----NSODSIEISKLDN-----KVAEYRDAYREYIAQMSQLEGPGSTTISGRSS 1355
Db      1279 FAKTNDPVASSLRCMCMDDRYDKLTQEOENEEVARS--KDIEVLEKPIYVOCYGLA 1336
QY      1356 PHSITYMGSSSGG-----SIHSLNEOE-----KKDSEPKPDGKRSFLKRGVDYDS 1405
Db      1337 P-----LTKGGOOLVFNFYSPFENRLPESIKIRDSOEP-CGRLSFLKE-----PKT 1382
QY      1406 SSGVSTNDASPLD-----PTEDEKSD--OSGSKLPGKK-----SSERSLFOPTLK 1452
Db      1383 TGLGQOTAVCNINTLPLAKKAKKADKROSFASLKRKYSLTEPSPSPQSPCERIDIR 1442
QY      1453 L-----KSGLRQYOKLPSDESESTE-----ESDNT-----PLLK-----DDRKAEG 1491
Db      1443 MAIVADHIGLSTWELARELNFVDEINQIRVENPNSLISQSPMLKKWTVTDGKNATGDA 1502
QY      1492 -----KYBRV-----KSPHSAP-PI 1507
Db      1503 LFSVLTKINRIDIVTLLEGPIFYGNISGTRSFADENNVFHDVYDQWONETPSGLSPA 1562
QY      1508 RPFIRAKEXLSDALLDKKSDSGVRS-----ESSPNHSLHNEVADDSOLEK 1555
Db      1563 QA-----RRLTGGLDRDDSDQARDSTTSLTEBPCKIEANGNHTA--EYIIPA--K 1612
QY      1556 ANLIELEDSSHSGKRG-----PHSLGLODPIIARMSICSEDKSPSBCS--LIASS 1606
Db      1613 ACPYPESQNDIGKOSIKENLKPKTHGCGRTPEV--SPLTAYOKSLEFYSKLYVIDA 1668
QY      1607 PEENPACOKAYNLARTSTYTLNNSAPARANONPEDEMGICRETSOYLIRPSSSPPT 1666
Db      1669 PKPCVPVGMK--KMTRTAD-----GKARLNOEERG-----STRSEPK 1705
QY      1667 TIQENELKSMTHK 1679
Db      1706 --QGEQYKVKTKK 1716

```

---

```

RESULT 14
ID 013768 PRELIMINARY: PRT: 1719 AA.
AC 013768:
DT 01-NOV-1996 (ITEMBLREL. 01, Created)
DT 01-NOV-1996 (ITEMBLREL. 01, Last sequence update)

```



```

Db      1476 RGEIYNNMLEGSGRGSRNLKPD-----RRHTDRDYSLSPSONMNGHORGARITHSPTV 1527
QY      1418 DPTREDEKS-----DQSGS-----KLBGKSSRSST.FOTDLKLSGSL 1458
Db      1528 SOVTRSDRIODMDADSIYSYLODAAGSGWQEEVYGGPPHFGQSTMTREGLR-PGSGQ 1586
QY      1459 RYOKLPSDESDSGTEESDNTPLLDKDKRAEGKVERYPKSPHSAEPITFTIKAEYLS 1518
Db      1587 EYEKLVSVSEHTWTEQPEASSQADRROQGOGEVY-----QKAKNTFYOVYOGNEFON 1642
QY      1519 ---DALLDKRDSSSG-----VRSESSPPNLSL--HNEVADDSOLEKANLIELE 1562
Db      1643 IPGEQVVEEOPFTDEOGNIVTKIIRKVVROIDLSSADAAOEHEEVEEGPLEDPSLEVD 1702
QY      1563 DD---SHS 1567
Db      1703 IDYFMKHS 1710

RESULT 15
ID      012955 PRELIMINARY; PRT: 4377 AA.
AC      012955:
DT      01-NOV-1996 (TREMblrel. 01, Created)
DT      01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Ankyrin G.
GN      ANK-3.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX      NCBI_TaxID=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      TISSUE-BRAIN STEM;
RA      MEDLINE=95138209; PubMed=7836469;
RX      Kordell E., Lambert S., Bennett V.;
RT      "Ankyrin, A new ankyrin gene with neural-specific isoforms localized
RT      at the axonal initial segment and node of Ranvier.";
RL      J. Biol. Chem. 270:2352-2359(1995).
RN      (2)
RP      SEQUENCE FROM N.A.
RC      TISSUE-BRAIN STEM;
RA      Carpenter S.S.;
RL      Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR      HSSP; P55273; 1B18.
DR      InterPro: IPR002110: ANK.
DR      InterPro: IPR000488: Death.
DR      InterPro: IPR000906; 205.
DR      Pfam: PF00023; ank; 24.
DR      Pfam: PF00531; death; 1.
DR      PRINTS: PR01415; ANKYRIN.
DR      SMART: SM00248; ANK; 21.
DR      SMART: SM00218; ZUS; 1.
DR      SMART: SM00005; DEATH; 1.
DR      PROSITE: PS50088; ANK_REPEAT; 21.
DR      PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR      PROSITE: PS50017; DEATH_DOMAIN; 1.
KW      ANK repeat; Repeat.
SQ      SEQUENCE 4377 AA; 480395 MW; F42379E55768B684 CRC64;

Query Match      5.9%; Score 521; DB 4; Length 4377;
Best Local Similarity 20.6%; Pred. No. 5,1e-23;
Matches 427; Conservative 284; Mismatches 695; Indels 668; Gaps 87;
QY      14 EENIIPALAKLKCKCDVDERNCGOTPLMIAAGNLEITVELKINGANGCNLEDDNMT 73
Db      83 KEHVEVSELLQREANVDATKKGNTALHIASLAGOAEVYKLVYNGANVANAQNGFT 142
QY      74 ALISAKEGHVHVEELKCGVNLERDMCGWALMAGCYKKGRTDVELLL----- 124

```

```

Db      143 PLVMAQENLEVEYKFLLDNGASQSLATEDGFTPLVAVLQOGHDVVSLLENDTKGVR 202
QY      125 -----SHGANSPTVL----- 135
Db      203 LPALHIAARKDPTKAALLLQNDNMNADVESKSGFTPLHIAAHGNINVTLLNRAAVD 262
QY      136 ---QYSYPLTMAAGSHADIVALLQNGAKVCSKYGTPPLVMARKG----- 182
Db      263 FTARNDTPLHVAASKRGNAVMKLLDRGAKIDAKTRDGLPLHCGARGSHQGVYBMLD 322
QY      183 -----HLECVKILAMGADVOEGANSWTALIVAKGSGTOS 219
Db      323 RAAPILSKTKNGSLPLMATQGDHLCYQILLQHNVPVDVYNDYLTPLHVAHCGHYAV 382
QY      220 KEILKRNPNVNLTDKGDNTALMASKE----- 247
Db      383 AKVLLDKKAPNAKALNGFTPLHACKKRIKYMELLKHGASIQAVTESGLPIHVAAF 442
QY      248 -GHTEIYODLDAGTYVINIPDRSGDVLIGAVGSHVEIYRALLQKADIDRGODNKTA 306
Db      443 MGVNIVISQLMHGHGASPNNTNNGETALHMAARSGOAEVRYLYODGAQVEAKADDQTP 502
QY      307 LYMAVEKGNATVYRDILO--CNPD-----EICT 333
Db      503 LHISARLGRADIVQOLLOQASPNNAFTSGYTPPLHSAREGHEDVAFLIDHGASISTT 562
QY      334 KGEFTPLIKATKRNIEVELLDKGAQVSAVDKGDTPPLHAIIRGRSKLAELLRNPK 393
Db      563 KKGFTPLHVAAKGKLEAVANLLLOKSASPDAGKSLGPLHVAAHYDNGKVALLLL--D 619
QY      394 DGLTLRPMKAGTEPNIDC--SHOKSLQI--FGAR-----HLSPTEDGDM 438
Db      620 QGASPIAANKGYTPPLHIAKKNMQDIATLLETLEGADANAVTROGIAVHLAAQGHVDM 679
QY      439 LGYDLYSALADLISEPTMQP-----PCVGLY 466
Db      680 VSLILGRNANVNLNKSNGTLPLHAAQEDRVAVVAEVLVQGAHVDAQTKMGTPPLHVGCH 739
QY      467 AOMGSK--SELIK-KLEDEMTFAQOIEPFPFSWILVPLTLICG--LCLLPAF 519
Db      740 --YGNIKIVNPLLOHSAKVNARTKNG--YTPHQAQ-----GQHTHINVLQON 785
QY      520 TVHPN-----LGIAVSLFALLYTFIVIFGGRGESNMMAVLSRLARH-I 569
Db      786 NSPNRLVYNGVNTALGIARLGIYSVDLKIYV-----EETMTTITTEKHM 834
QY      570 GYIELLLKIMFVNPEL-----PEQTKALPVRFETDYNRLSY-GEFTSLA----- 616
Db      835 NVEPTWNEVILMDSDEVRKANAP-----MLSDGEYISDVEGEDAWGTDXYL 884
QY      617 --EMATLSDACEREGFLATRL-----FRYFKED--TQKKKKWKTCCLPSEVILF 666
Db      885 GPQDLKEIGDSDLPAEGYMGFSIGARSASLSRFSDDRSYTLNRSSYAR-----DSMMIEL 940
QY      667 IIGCLISGITLLAIFRVPD-KHLT-----VNAVLSIASVGLAVLNCRMTWQYLD 717
Db      941 IYPSKQKHLLTFRPDSLSLRHYSWAADPLDWNVYSSPIHSGFLVSTWVDKRG----- 994
QY      718 SLNSQRKRLHNAASKLRLKLGEGFMKYL--KCEVELMARMAKTIDSFQONQRLVVI 773
Db      995 ---GSMRGSRHHG-----MRIIIPRKTAP-----TRICRLV-- 1025
QY      774 IGDLDACBDQVLYQMLDVIYRVLFSGPFIAPSPHIITAIINONLSYLRDSINGHD 833
Db      1026 -----KHKLIAN-----PPH-----GERGJSSRLVEMGPAPAQ 1055
QY      834 YMRN-IYHLPYVLSRGLSNARKFLVTSATNGDVCSDTGTQEDADRRVONSIGEXTK 892
Db      1056 FLGPVIVELPHRGSRG--KERELIVLSENCE-----YKHEQFD--SKNE--DLTE 1102
QY      893 L--GSKTALNRDQYRRQOMORTITROMSFDLTLLVTEWFSDISPQTMRLINIVSVT 950

```

DB 1103 LINGDELDSPDELKRRICRITTKDP-----POYFAVVS--RIKQESNOIGPE 1150  
QY 951 GRLLRAN-----QISFNMWDLASWINLEQMPYRTSWLILYEETEGIPDOMTLKTIYER 1005  
DB 1151 GGILSSTVPLVQASPEGALTKRIRVGLQ-----AQPVPEIYKILGNK 1196  
QY 1006 ISKNIPPTKDYEPLELDGDRNEFEVLSKTPYLVAROVKFLPCTVINLDPKIRELIAD 1065  
DB 1197 AT-----FSPITVEPRRRKFKPITMTIPV----- 1222  
QY 1066 VRAAREQISIGLAVPLPLEGPPRAPSGYSCP--PSVCSSTFNGPAGVVSPOPHS 1123  
DB 1223 -----PPSGEGVSNKGDTTPNLKLCST-----GGTSP----- 1254  
QY 1124 SYISGMTGPQHPFYNRGSGPAPGVLLIN-----SLAVD-----AVCEKLOLEGLDQSM 1173  
DB 1255 AQMEDITGTT-----PLTFIKDCVSFTTWSARFWLADCHQVLETVGLATOL 1301  
QY 1174 -----LPQYCTTIKKANINGRVLQOCNI-----DELKK--EMMNFQDMHLFRSTVLEM 1220  
DB 1302 YRELICVYMAKFVYFAPKANDPVESSLRCFCMTDDKVDKTLQOENFE-----VA 1352  
QY 1221 RNASHVVPEDDPRELSESSSGPAPHGEPARA-----SHNELPHT-----ELSSQTPYT 1269  
DB 1353 RSKDIEVLEKPIYV--DCYGNLAPLTKGGQOLVFENFYSFKENRLPFISIKINDTSQEPCCR 1411  
QY 1270 LNFSEELNTLGLDEGAPRHSNLSWOSQTRKTPSLSLNSODSSIETSKLTDKQVQAEYRD 1329  
DB 1412 LSFLEKERTTGLPOTAVCNINITLPAHKETES-----DQDDEIE--KTDROQSPASL 1463  
QY 1330 AYRE---YIAQMSOLEGGPGSTTISGRSPHSTY-----YMGSSSGSISHSNLE 1376  
DB 1464 ALRKRYSTLTERPGMTERSTGAT-----RSLP--TYSYKPFSTRPYQSMTAPITV----- 1513  
QY 1377 QEKGDSEPKPDGGRKSLMRGVDYIDYSSSGVSTNDASPLDPI---TEDEKSDQSGS 1432  
DB 1514 -----PGPAKSGF-----TSLSSSSSNTPSASPLKISWSVSTPSPIKSTLGAS 1556  
QY 1433 KLIPGKKSERSSLFQTLKLGSLR---YOKLPSDEDESGT-----EESDNPL----- 1480  
DB 1557 TTSSVKSISDVASPIRS--LRTWSSPIKTVSOSOPYNIQVSSGTLARAPAVTEATPLKGLA 1615  
QY 1481 -----LKDDKDKRAEGKVER-----VPKSP-----HSAEPIRFP--KAKEYLSDAL 1521  
DB 1616 SMTSSSRTPYTGTGSLERSSITMTPPASPKSNINMYSSSLPFKSLITSAAPLISPL 1675  
QY 1522 -----LDKDDSDSGVRSSESPNHSL--HNEVA---DDSOLEKAN---LIE-- 1560  
DB 1676 KSVSPVKSRYDVVISAKITMASSLSISPVKQMPGHAVALYNGSISPLKYASSSTLNGC 1735  
QY 1561 -----LEDDSHGKRGIPHSLSGLDDPIIARMSICSEDKSPSECSLIASSPEENWPACQ 1615  
DB 1736 KATATLOEKISSATNSVSSVSAADYEVKFTITAMPFSPLR--SYSAAP----- 1786  
QY 1616 KAYNLNRTPTSTVTLNNSAPANRANQNDMEGIRETSOVLIRPSSS-----PNFTIQN 1670  
DB 1787 SAFQSLRTPPSASALYTSIGSSISATTS-----SVTSIITVPYSVVNVLPBPAKLKL 1839  
QY 1671 ENLKSMTHKRSQRS--YTRLKSDPPELHAAAS 1702  
DB 1840 PDSNSFTKSAALLSPIKTLTETHPQPHFSRTS 1873

Search completed: July 1, 2003, 14:40:31  
Job time : 112 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 55.5 Seconds

(without alignments)  
4117.564 Million cell updates/sec

Title: US-10-021-571-2

Sequence: 1 MSVLISQSVINVEENIPV.....ELHAASSTSGTGERESTL 1715

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Archived: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8639.5	97.6	1762	23	AAU96841	Rat kidneys220 prot
2	8526	96.3	1763	23	AAU96824	Human polypeptide
3	8266	93.4	1715	22	AAU96893	Human polypeptide
4	8266	93.4	1715	22	AAU96840	Human polypeptide
5	8266	93.4	1715	22	AAU96840	Human polypeptide
6	2961.5	33.5	705	22	AAU96840	Human polypeptide
7	2787.5	31.4	551	22	AAU96840	Human polypeptide
8	2408	27.2	543	22	AAU96840	Human polypeptide
9	2376.5	26.8	1498	22	AAU96840	Human polypeptide
10	2260	25.5	513	22	AAU96840	Human polypeptide

11	2150	24.3	624	22	ABG08697	Novel human diago
12	614	6.9	120	22	AAU40811	Human polypeptide
13	599.5	6.8	129	22	AAU94785	Human protein sequ
14	549	6.2	109	21	AAU56277	Human secreted pro
15	538.5	6.1	1724	22	AAU20513	Protein of App rel
16	534.5	6.0	1872	22	AAU9160	Human protein, SEQ
17	526	5.9	4274	22	ABG00972	Novel human diago
18	521	5.9	4386	22	ABG07375	Novel human diago
19	521	5.9	4397	22	ABG21944	Novel human diago
20	504	5.7	2443	22	ABG60521	Drosophila melanog
21	500.5	5.7	187	22	AAU40779	Human polypeptide
22	492.5	5.6	784	22	ABG62206	Drosophila melanog
23	471	5.3	2119	22	ABG64823	Drosophila melanog
24	467.5	5.3	1549	22	ABG58328	Drosophila melanog
25	467.5	5.3	1549	22	ABG67412	Drosophila melanog
26	466.5	5.3	1745	19	AAU70608	Full length ankyl
27	466.5	5.3	1745	19	AAU70608	D. limulitis ankyl
28	466.5	5.3	1745	19	AAU70608	D. limulitis ankyl
29	466.5	5.3	1745	23	AAU1589	Ankyrin protein se
30	459	5.2	743	23	ABG97453	Novel human protei
31	450	5.1	1133	22	ABG08071	Novel human diago
32	449.5	5.1	1088	22	ABG62429	Human kidney ankyr
33	449.5	5.1	1097	22	ABG21941	Novel human diago
34	448	5.1	378	22	AAU64963	Human secreted pro
35	441	5.0	435	22	AAU66710	Human cell growth
36	441	5.0	435	22	AAU93879	Human protein sequ
37	441	5.0	627	23	AAU17136	Human cancer cell
38	438	4.9	740	22	AAU01030	Human cytoskeleton
39	438	4.9	1569	22	AAU74360	Human diago
40	434	4.9	1377	22	ABG08072	Novel human diago
41	432	4.9	1267	22	ABG22790	Novel human diago
42	431.5	4.9	1053	22	AAU28174	Novel human secret
43	429	4.8	378	22	AAU64962	Gene 24 human secr
44	423.5	4.8	705	22	AAU20496	Human secreted pro
45	421.5	4.8	2441	22	ABG62231	Drosophila melanog

# ALIGNMENTS

RESULT 1  
AAU96841  
ID AAU96841 standard; Protein: 1762 AA.  
XX  
AC AAU96841;  
XX  
DT 30-JUL-2002 (first entry)  
XX  
DE Rat kidneys220 protein from Genbank AAU96841.  
XX  
DE Rat kidneys220 protein from Genbank AAU96841.  
XX  
KW Rat kidneys220; kinase D interacting substrate of 22kDa; cytosolic;  
KW neuroprotective; gene therapy; protein kinase D; PDK; cancer; rat;  
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.  
XX  
OS Rattus sp.  
XX  
PN WO20020786-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 06-SEP-2001; 2001MO-GB03977.  
XX  
PR 06-SEP-2000; 2000US-230449P.  
XX  
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.  
XX  
PI Schiavo G, Iglesias T;  
XX  
DR WPI; 2002-371879/40.  
XX  
PT Polypeptide kinase D interacting substrate of 220 kDa and  
PT polynucleotide; useful for identifying modulators useful in treating  
PT cancer and neurodegenerative diseases

XX Disclosure; Fig 18; 228pp; English.

CC The invention relates to polypeptide comprising a kinase D interacting  
CC substrate of 220 kDa (Kidins220) from rat and human or their  
CC fragment, variant or fusion provided that the protein is not the  
CC polypeptide encoded by Genbank Accession No. AB033076. Also included are  
CC a polynucleotide encoding Kidins220 provided that polynucleotide is  
CC not any one of the clones corresponding to the 61 Genbank Accession Nos.  
CC given in the specification, an expression/replicable vector comprising  
CC the polynucleotide; a recombinant host cell containing the polynucleotide  
CC or vector; an anti-Kidins220 antibody (used in the preparation of  
CC Kidins220); an agent (AI) which modulates activity of Kidins220;  
CC protein kinase D (PKD) or interaction between PKD and Kidins220;  
CC a mutant animal transgenic for Kidins220; the use of an agent capable of  
CC detecting the expression of Kidins220 gene in manufacture of a diagnostic  
CC reagent for diagnosing or prognosing cancer or for monitoring the  
CC progression of cancer in a patient; a compound comprising a moiety that  
CC selectively binds to Kidins220 or its variant and another moiety; and a  
CC nucleic acid molecule encoding the compound. The antibody is useful for  
CC modulating the activity of Kidins220, where the antibody prevents a  
CC region of Kidins220 interacting from another protein, or reduces the  
CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The  
CC Kidins220 polynucleotide is useful for identifying an agent which  
CC modulates the promoter activity of the polynucleotide, and Kidins220 is  
CC useful for identifying an agent which modulates activity of Kidins220,  
CC protein kinase D (PKD) and the interaction between PKD with Kidins220.  
CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene  
CC therapy or for diagnosing or monitoring progression of cancer in a  
CC patient. The agent is useful in manufacture of medicament, for use in  
CC treatment of neurodegenerative disease, and the agent which inhibits  
CC function of Kidins220 gene or its product is useful for treating cancer  
CC which is glioblastoma multiforme or prostate cancer. The compound is  
CC useful for imaging cancer in an individual, and for diagnosing or  
CC prognosing, and also for treating an individual. The present  
CC sequence represents Rat Kidins220 protein from Genbank accession number  
CC AA035185.

XX Sequence 1762 AA;

Query Match 97.6%; Score 8639.5; DB 23; Length 1762;

Best Local Similarity 95.8%; Pred. No. 0; Mismatches 15; Indels 53; Gaps 5;

Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;

QY 1 MSVLISGVINVEENIPALKALLECKDVERNEGQPTPLMAEGVNEIYKELLN 60  
Db 1 MSVLISGVINVEENIPALKALLECKDVERNEGQPTPLMAEGVNEIYKELLN 60

QY 61 GANCNLEDDNMWTLALISAKSGHITHYELLSKASLEHRDMGWTALMACTYKGRPDVY 120  
Db 61 GANCNLEDDNMWTLALISAKSGHITHYELLSKASLEHRDMGWTALMACTYKGRPDVY 120

QY 121 ELLLSHGANSVPTGLQTSYPTIWAAGRGHADIYHLLONGAKVNSDXTGTPPLWAAR 180  
Db 121 ELLLSHGANSVPTGLQTSYPTIWAAGRGHADIYHLLONGAKVNSDXTGTPPLWAAR 180

QY 121 ELLLSHGANSVPTGLQTSYPTIWAAGRGHADIYHLLONGAKVNSDXTGTPPLWAAR 179  
Db 121 ELLLSHGANSVPTGLQTSYPTIWAAGRGHADIYHLLONGAKVNSDXTGTPPLWAAR 179

QY 181 KGHLECYKHLIAGADVDEGANSMTALIAVKGQTSYKVELLRPNVNLTDKQNTA 240  
Db 181 KGHLECYKHLIAGADVDEGANSMTALIAVKGQTSYKVELLRPNVNLTDKQNTA 240

QY 180 KGHLECYKHLIAGADVDEGANSMTALIAVKGQTSYKVELLRPNVNLTDKQNTA 239  
Db 180 KGHLECYKHLIAGADVDEGANSMTALIAVKGQTSYKVELLRPNVNLTDKQNTA 239

QY 241 LMTASRGHTEIYODLLDAGTYVNIIPRSGDYVLIGVNRGHEIYVALLQKADIDIRG 300  
Db 241 LMTASRGHTEIYODLLDAGTYVNIIPRSGDYVLIGVNRGHEIYVALLQKADIDIRG 300

QY 240 LMTASRGHTEIYODLLDAGTYVNIIPRSGDYVLIGVNRGHEIYVALLQKADIDIRG 299  
Db 240 LMTASRGHTEIYODLLDAGTYVNIIPRSGDYVLIGVNRGHEIYVALLQKADIDIRG 299

QY 301 QDNKTALYMAVEKGNATVNDILOCNPDTEICTKDGTEPLIKATKMNIEVEVLLDKGA 360  
Db 301 QDNKTALYMAVEKGNATVNDILOCNPDTEICTKDGTEPLIKATKMNIEVEVLLDKGA 360

QY 300 QDNKTALYMAVEKGNATVNDILOCNPDTEICTKDGTEPLIKATKMNIEVEVLLDKGA 359  
Db 300 QDNKTALYMAVEKGNATVNDILOCNPDTEICTKDGTEPLIKATKMNIEVEVLLDKGA 359

QY 361 KVSAYDKKGTPLHVAIRGSRRLAELLRNPKDGRLLYRNKAGETPNYNDISHSKSLI 420  
Db 361 KVSAYDKKGTPLHVAIRGSRRLAELLRNPKDGRLLYRNKAGETPNYNDISHSKSLI 420

QY 360 KVSAYDKKGTPLHVAIRGSRRLAELLRNPKDGRLLYRNKAGETPNYNDISHSKSLI 419  
Db 360 KVSAYDKKGTPLHVAIRGSRRLAELLRNPKDGRLLYRNKAGETPNYNDISHSKSLI 419

QY 421 TQIFGARHLSPTETDGMGLGYDLYSSALMDIILSEPIMQPTICVGLYANQSGSFLKLKL 480  
Db 421 TQIFGARHLSPTETDGMGLGYDLYSSALMDIILSEPIMQPTICVGLYANQSGSFLKLKL 480

Db 420 TQIFGARHLSPTETDGMGLGYDLYSSALMDIILSEPIMQPTICVGLYANQSGSFLKLKL 479

QY 481 EDEKKTFRAGQOTEPLEQFSWLIYVTLTLCGGGLVFAFPVNDNLAIASLSFALIYTF 540  
Db 480 EDEKKTFRAGQOTEPLEQFSWLIYVTLTLCGGGLVFAFPVNDNLAIASLSFALIYTF 539

QY 541 FIYIYGGRRGEGSWMMANALSTRLARHIGYLELFKLMFVNDPELPEOTYALPVRLF 600  
Db 540 FIYIYGGRRGEGSWMMANALSTRLARHIGYLELFKLMFVNDPELPEOTYALPVRLF 599

QY 601 TDNRRLSVGGEISLAEIATLSDACEREGFATLRFVFRRESQKKKKKTCCLPS 660  
Db 600 TDNRRLSVGGEISLAEIATLSDACEREGFATLRFVFRRESQKKKKKTCCLPS 659

QY 661 FVIFLEFVIGCIAGITLLAIFRVPKHLTVNALISIASVGLAFVLCRTMVOYLDLIL 720  
Db 660 FVIFLEFVIGCIAGITLLAIFRVPKHLTVNALISIASVGLAFVLCRTMVOYLDLIL 719

QY 721 NSQKRLHSAASKLHKLKSGEFKVLKCEVELMARAKTIDFTQNTQRLVYIDGLDAC 779  
Db 720 NSQKRLHSAASKLHKLKSGEFKVLKCEVELMARAKTIDFTQNTQRLVYIDGLDAC 779

QY 781 EODKVLQMDTVAVLFSKGEPIAFASDPHIITKAINQMLNSYLRDSNINGDYKRNIVH 840  
Db 780 EODKVLQMDTVAVLFSKGEPIAFASDPHIITKAINQMLNSYLRDSNINGDYKRNIVH 839

QY 841 LPVFLNSRGLSNARKLTVSATNGDITGSDTGTQEDTDRLVRSQNSLGEHTKSGTALN 900  
Db 840 LPVFLNSRGLSNARKLTVSATNGDITGSDTGTQEDTDRLVRSQNSLGEHTKSGTALN 899

QY 901 RRTYRRRQORITQMSFDTLKLTVTDWSDISPOMRRLNIVSTGLTALANOT 960  
Db 900 RRTYRRRQORITQMSFDTLKLTVTDWSDISPOMRRLNIVSTGLTALANOT 959

QY 961 FNMRLASWINTEQWPRYSWILYLEETEGLPDQMTLKYERISKNIPTTKDVEPLL 1020  
Db 960 FNMRLASWINTEQWPRYSWILYLEETEGLPDQMTLKYERISKNIPTTKDVEPLL 1019

QY 1021 EIDGDIRNEFVLLSTSTPVLVADVKTFLPCVYNLDPKRELIAVRAAREINIGGLAY 1080  
Db 1020 EIDGDIRNEFVLLSTSTPVLVADVKTFLPCVYNLDPKRELIAVRAAREINIGGLAY 1079

QY 1081 PPLPLEHGPFRPSGSGQSPASVCSASFNGPPGGVSPQPSSTYSGLSGQHPFYNR 1139  
Db 1080 PPLPLEHGPFRPSGSGQSPASVCSASFNGPPGGVSPQPSSTYSGLSGQHPFYNR 1139

QY 1140 -----AAPTATSSLLL 1151  
Db 1140 -----AAPTATSSLLL 1151

QY 1152 SSMYVNVCEKLRQTEGLDQNNMPOCTTIKANKNGVRLSCNIDELKKEAMFQWH 1211  
Db 1200 SSMYVNVCEKLRQTEGLDQNNMPOCTTIKANKNGVRLSCNIDELKKEAMFQWH 1209

QY 1212 LFRSMVLEKRSVESQVPEDPFLNENSSAPVPHESARRSHTLPLTELSSOPTYLN 1271  
Db 1260 LFRSMVLEKRSVESQVPEDPFLNENSSAPVPHESARRSHTLPLTELSSOPTYLN 1319

QY 1272 FSFEELNTLGLDEGAPRHSNLSWSQTRTPSLSLNSQDSSIETSLTRVQAEYRDAY 1331  
Db 1320 FSFEELNTLGLDEGAPRHSNLSWSQTRTPSLSLNSQDSSIETSLTRVQAEYRDAY 1379

QY 1332 REYIAMSQLEGGTSSITSGSSPHSTYVYIGOSSGGSISHTLEOEGKGEKLOEGR 1391  
Db 1380 REYIAMSQLEGGTSSITSGSSPHSTYVYIGOSSGGSISHTLEOEGKGEKLOEGR 1439

QY 1392 KSFLLKRGVDYIDSSSGVSTNEASPLDITTEDEKSDSGSKLPGKSSERPLFOTDL 1451  
Db 1440 KSFLLKRGVDYIDSSSGVSTNEASPLDITTEDEKSDSGSKLPGKSSERPLFOTDL 1499

QY 1452 KIKGGGLKOKLPSDESSGTGRVQITPHCSKMTKRLAKK -ORECASPOEHSAPERT 1510  
Db 1452 KIKGGGLKOKLPSDESSGTGRVQITPHCSKMTKRLAKK -ORECASPOEHSAPERT 1510

Db 1500 KIKGGLRYOKLPDEDESGTEESDNTF-LTKDKDKAKGKAEKRYCKSP-EHSAEPIRT 1557  
 QY 1511 FIKAKXYLSDALLDKRSDSGVRSNESPNSHNEADDSQLEKANIIELEDGSHGK 1570  
 Db 1558 FIKAKXYLSDALLDKRSDSGVRSNESPNSHNEADDSQLEKANIIELEDGSHGK 1617  
 QY 1571 RGMPHSLGLDPIIIRMSICSEDKSPSECSLIASSPESPACOKAYNLNRTPTSTYL 1630  
 Db 1618 RGMPHSLGLDPIIIRMSICSEDKSPSECSLIASSPESPACOKAYNLNRTPTSTYL 1677  
 QY 1631 NNNTAPTNANONFDEIEGIRETSOYLIRPGSPNPATAVONENLKSMAHRSQSSYTRL 1690  
 Db 1678 NNNTAPTNANONFDEIEGIRETSOYLIRPGSPNPATAVONENLKSMAHRSQSSYTRL 1737  
 QY 1691 SKDASELHAASSESTGFEGERESIL 1715  
 Db 1738 SKDASELHAASSESTGFEGERESIL 1762  
 RESULT 2  
 ID AA080244  
 AA080244 standard; Protein; 1763 AA.  
 AA080244;  
 30-JUL-2002 (first entry)  
 DE Rat Kidins220 protein.  
 KW Kidins220; kinase D interacting substrate of 22kDa; cytosolic;  
 KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;  
 KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.  
 OS Rattus sp.  
 PN WO200220786-A2.  
 PD 14-MAR-2002.  
 PF 06-SEP-2001: 2001MO-GB03977.  
 PR 06-SEP-2000: 2000US-230449P.  
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 PI Schiavo G, Iglesias T;  
 WP1; 2002-371879/40.  
 N-PSDB; ABR51214.  
 PT Polypeptide kinase D interacting substrate of 220 kDa and  
 PT polynucleotide; useful for identifying modulators useful in treating  
 cancer and neurodegenerative diseases  
 PS Claim 33; Fig 1; 228pp; English.  
 CC The invention relates to polypeptide comprising a kinase D interacting  
 CC substrate of 220 kDa (Kidins220) from rat and human or the  
 CC fragment, variant or fusion provided that the protein is not the  
 CC polypeptide encoded by Genbank Accession No. AB033076. Also included are  
 CC a polynucleotide encoding Kidins220 provided that polynucleotide is  
 CC not any one of the clones corresponding to the 61 Genbank Accession Nos.  
 CC given in the specification; an expression/replacable vector comprising  
 CC the polynucleotide; a recombinant host cell containing the polynucleotide  
 CC or vector; an anti-Kidins220 antibody (used in the preparation of  
 CC Kidins220); an agent (AI) which modulates activity of Kidins220;  
 CC protein kinase D (PKD) or interaction between PKD and Kidins220;  
 CC a mutant animal transgenic for Kidins220; the use of an agent capable of  
 CC detecting the expression of Kidins220 gene in manufacture of a diagnostic  
 CC reagent for diagnosing or prognosing cancer or for monitoring the  
 CC progression of cancer in a patient; a compound comprising a moiety that  
 CC selectively binds to Kidins220 or its variant and another moiety; and a  
 CC nucleic acid molecule encoding the compound. The antibody is useful for

CC modulating the activity of Kidins220, where the antibody prevents a  
 CC region of Kidins220 interacting from another protein, or reduces the  
 CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The  
 CC Kidins220 polynucleotide is useful for identifying an agent which  
 CC modulates the promoter activity of the polynucleotide, and Kidins220 is  
 CC useful for identifying an agent which modulates activity of Kidins220,  
 CC protein kinase D (PKD) and the interaction between PKD with Kidins220.  
 CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene  
 CC therapy or for diagnosing or monitoring progression of cancer in a  
 CC patient. The agent is useful in manufacture of medicament, for use in  
 CC treatment of neurodegenerative disease, and the agent which inhibits  
 CC function of Kidins220 gene or its product is useful for treating cancer  
 CC which is glioblastoma multiforme or prostate cancer. The compound is  
 CC useful for imaging cancer in an individual, and for diagnosing or  
 CC prognosing, and also for treating an individual. The present  
 CC sequence represents Rat Kidins220.  
 CC  
 SQ Sequence 1763 AA;  
 Query Match 96.3%; Score 8526; DB 23; Length 1763;  
 Best Local Similarity 94.7%; Pred. No. 0;  
 Matches 1671; Conservative 10; Mismatches 32; Indels 52; Gaps 6;

QY 1 MSVLISQSVINVEENIPALKALLKCKDVERNECGQTPMLAAGNVEIKELLN 60  
 Db 1 MSVLISQSVINVEENIPALKALLKCKDVERNECGQTPMLAAGNVEIKELLN 60  
 QY 61 GANCLLELDLWMTALISKKEGHIHYBELLSKASLEHRMGWGTALMAACRYGRDYY 120  
 Db 61 GANCLLELDLWMTALISKKEGHIHYBELLSKASLEHRMGWGTALMAACRYGRDYY 120  
 QY 121 ELLSHGANPVTGLOYSVYPIIWAAGRGHADYIHLHLONGAKVNCSDKGTTPVLAAR 180  
 Db 121 ELLSHGANPVTGLOYSVYPIIWAAGRGHADYIHLHLONGAKVNCSDKGTTPVLAAR 180  
 QY 181 KGLEECVYKHLAMGADVDOEGANSMTALIVAAGGYQSVKEILKRNPNVLDKDGNTA 240  
 Db 181 KGLEECVYKHLAMGADVDOEGANSMTALIVAAGGYQSVKEILKRNPNVLDKDGNTA 240  
 QY 241 LMIAKKEGHIETVDLDAGTYVINIPRSGTYVIGAVRGHVEIVALLQKADIDIRG 300  
 Db 241 LMIAKKEGHIETVDLDAGTYVINIPRSGTYVIGAVRGHVEIVALLQKADIDIRG 300  
 QY 301 ODNKATLWYAEKGNATVWRDILQCNPTETCTKDGETPLIKATKMRNIEVEELLDKGA 360  
 Db 301 ODNKATLWYAEKGNATVWRDILQCNPTETCTKDGETPLIKATKMRNIEVEELLDKGA 360  
 QY 361 RYSAVDDKGGDPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPYNIDCSHOKSIL 420  
 Db 361 RYSAVDDKGGDPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPYNIDCSHOKSIL 420  
 QY 421 TOIFGARHLSPETDGMGYDLVSSALADILSEPTMQPPTCVGLYQWMSGKSFLLKTL 480  
 Db 421 TOIFGARHLSPETDGMGYDLVSSALADILSEPTMQPPTCVGLYQWMSGKSFLLKTL 480  
 QY 481 EDEKTFAGQOQTEPLFOFSMLIVFTLLICGGLGVFAFVDTMLATAISLFFALIYIF 540  
 Db 481 EDEKTFAGQOQTEPLFOFSMLIVFTLLICGGLGVFAFVDTMLATAISLFFALIYIF 540  
 QY 541 FIVIFEGRRGEESNMWAMALSTRLARHIGYLELLFKIMFVNPELPBQTTKALPVRLF 600  
 Db 541 FIVIFEGRRGEESNMWAMALSTRLARHIGYLELLFKIMFVNPELPBQTTKALPVRLF 600  
 QY 601 TDYNLSSVGGTSLAEIMATLSADCEEFGLATRLFRVPRTEESGKKKKWKTCCPLS 660  
 Db 601 TDYNLSSVGGTSLAEIMATLSADCEEFGLATRLFRVPRTEESGKKKKWKTCCPLS 660  
 QY 661 FVIFLFIYGCITAGITLLAIFRVPKHLTVNAIISIVAGLAFVLCRTMVOVLDLIL 720  
 Db 661 FVIFLFIYGCITAGITLLAIFRVPKHLTVNAIISIVAGLAFVLCRTMVOVLDLIL 720  
 QY 721 NSQRRRLSAAKHLIKLSEGPMLVKCEVELMAAMAKTIDSPONOTRLVVIIDGLDAC 780  
 Db 721 NSQRRRLSAAKHLIKLSEGPMLVKCEVELMAAMAKTIDSPONOTRLVVIIDGLDAC 780

```

Db      721 NSORRKLHSAASKLHKLKSEGFMKVYLKCEVELAMARAKTIDSFQNTQVYIIDGLDAC 780
QY      781 EDDKVLQMLDVTYVLFESKPEFIAIPASDPHIIIAKINONLSVLRDSNINGHDIYRIVH 840
Db      781 EDDKVLQMLDVTYVLFESKPEFIAIPASDPHIIIAKINONLSVLRDSNINGHDIYRIVH 840
QY      841 LPEVFNLSRGLSNARKFLVTSATNGDITGDTGTGEDDRVRVSONSLGEMFLGSKTALN 900
Db      841 LPEVFNLSRGLSNARKFLVTSATNGDITGDTGTGEDDRVRVSONSLGEMFLGSKTALN 900
QY      901 RRDYRRRQMQRTITRQMSFDLTKLLVTEWEDSDISPOIMRLNIVSVTGRLLRANOIT 960
Db      901 RRDYRRRQMQRTITRQMSFDLTKLLVTEWEDSDISPOIMRLNIVSVTGRLLRANOIT 960
QY      961 FMMDLASWINLTEMQPYRTSMWLLILEETEGLPQOMLTKTYEISKNIPPTKVEPL 1020
Db      961 FMMDLASWINLTEMQPYRTSMWLLILEETEGLPQOMLTKTYEISKNIPPTKVEPL 1020
QY      1021 EIDGDIRNEVEFLSSRTPLVARDVKTFLPCTVNLDPKLRITADYRARAREQINIGLAY 1080
Db      1021 EIDGDIRNEVEFLSSRTPLVARDVKTFLPCTVNLDPKLRITADYRARAREQINIGLAY 1080
QY      1081 PELPLHEGPPRPSSQSPASVCSASFNGPPGCVSPQPHSSVYSGLSGPHQFYNR- 1139
Db      1081 PELPLHEGPPRPSSQSPASVCSASFNGPPGCVSPQPHSSVYSGLSGPHQFYNR- 1139
QY      1140 -----AAVP-----ANGSSLLSSMTVD 1157
Db      1141 FPAPLYLPYPPYPGSQHLLISHSVAKALPRDQNNGLPCDSGFKRQGSRCRHLQVLT 1200
QY      1158 VV-----CEKLRQIEGLDQNMMPQCTTIKKANINGVLSOCNIDELKKEAMNFGDM 1211
Db      1201 EFNDGRCREKLRQIEGLDQNMMPQCTTIKKANINGVLSOCNIDELKKEAMNFGDM 1260
QY      1212 LERSVWLEMRSVESQVPEDPFLFNENSGAPVPHGESARRSHHELPLTELSQPTPTLN 1271
Db      1261 LERSVWLEMRSVESQVPEDPFLFNENSGAPVPHGESARRSHHELPLTELSQPTPTLN 1320
QY      1272 FSEFEELNTGLDEGAPRHSNLSMOSQTRTPSLSLNSODSSIEFSTKLTIDVOAEYRAY 1331
Db      1331 FSEFEELNTGLDEGAPRHSNLSMOSQTRTPSLSLNSODSSIEFSTKLTIDVOAEYRAY 1380
QY      1332 REYIAQMSQLEGSTISGRSPHSTYYIGQSSSGSISHTLEQERKGEELKQEDGR 1391
Db      1381 REYIAQMSQLEGSTISGRSPHSTYYIGQSSSGSISHTLEQERKGEELKQEDGR 1440
QY      1392 KSEFLMRGDDVIDYSSGVSTNEASPLDITEDEKSDSGSKLPGKSSSRPSPFQTDL 1451
Db      1441 KSEFLMRGDDVIDYSSGVSTNEASPLDITEDEKSDSGSKLPGKSSSRPSPFQTDL 1500
QY      1452 KLGKGLRYOKLPSDESGSTGRVQITPHCSKMITRKLAK-ORECASPOEHSAPERT 1510
Db      1501 KLGKGLRYOKLPSDESGSTGRVQITPHCSKMITRKLAK-ORECASPOEHSAPERT 1558
QY      1511 FIKAKYVLSDALDKKSDSGSVRSNESSPNHSLHNEAADSOLEKANLLELEDEGSHGK 1570
Db      1559 FIKAKYVLSDALDKKSDSGSVRSNESSPNHSLHNEAADSOLEKANLLELEDEGSHGK 1618
QY      1571 RGMPSLISGLDPIIARMSISSEKSPSECSLIASSPEEWPAQKRYNLRNPSTVTL 1630
Db      1619 RGMPSLISGLDPIIARMSISSEKSPSECSLIASSPEEWPAQKRYNLRNPSTVTL 1678
QY      1631 NNNTAPTRANQNFDEIGIRETSQVILRPSPNPZAVQENKLSMAHKRSQSSSYTRL 1690
Db      1679 NNNTAPTRANQNFDEIGIRETSQVILRPSPNPZAVQENKLSMAHKRSQSSSYTRL 1738
QY      1691 SKDASELHAASSESTGFEERESILI 1715
Db      1739 SKDASELHAASSESTGFEERESILI 1763

```

RESULT 3  
AAM38993

```

ID      AAM38993 standard; Protein; 1715 AA.
XX
AC      AAM38993;
XX
DT      22-OCT-2001 (first entry)
XX
DE      Human polypeptide SEQ ID NO 2138.
XX
KW      Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW      peripheral nervous system; neuropathy; central nervous system; CNS;
KW      Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW      amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW      chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW      leukaemia.
XX
OS      Homo sapiens.
XX
PN      MO20015312-A1.
XX
PD      26-JUL-2001.
XX
PE      26-DEC-2000; 2000MO-US34263.
XX
PR      21-JAN-2000; 2000US-0488725.
PR      25-APR-2000; 2000US-0552317.
PR      09-JUL-2000; 2000US-0598042.
PR      19-JUL-2000; 2000US-0620312.
PR      03-AUG-2000; 2000US-0653450.
PR      14-SEP-2000; 2000US-0662191.
PR      19-OCT-2000; 2000US-0693036.
PR      29-NOV-2000; 2000US-0727344.
XX
PA      (HYSE-) HYSED INC.
XX
PI      Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI      Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI      Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR      WPI: 2001-442253/47.
XX
N-PSDB: AAI58149.
XX
PT      Novel nucleic acids and polypeptides, useful for treating disorders
PT      such as central nervous system injuries -
XX
PS      Example 4; SEQ ID NO 2138; 10078bp; English.
XX
CC      The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC      the encoded polypeptides (AAM38642-AA42213) with nontropic,
CC      immunosuppressant and cytostatic activity. The polynucleotides are useful
CC      in gene therapy. A composition containing a polypeptide or polynucleotide
CC      of the invention may be used to treat diseases of the peripheral nervous
CC      system, such as peripheral nervous injuries, peripheral neuropathy and
CC      localized neuropathies and central nervous system diseases, such as
CC      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC      utilisation of the activities such as: Immune system suppression,
CC      Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC      assays for receptor activity, arthritis and inflammation, leukaemia and
CC      C.N.S disorders.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification.
XX
SQ      Sequence 1715 AA:
XX

```

Query Match 93.4%; Score 8266; DB 22; Length 1715;  
Best Local Similarity 92.7%; Pred. No. 0;  
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

```

QY      1 MSVLTISQSYINVEENIPALAKLEKCDVDERNECGTPTMLAAEGQNVIVVELLN 60
Db      1 MSVLTISQSYINVEENIPALAKLEKCDVDERNECGTPTMLAAEGQNVIVVELLN 60

```

QY 61 GANCNLEDDNMTALISASKEGHIIVEELLKSGSLERHDMGWTALMACYKGRPDVY 120  
 DB 61 GANCNLEDDNMTALISASKEGHIIVEELLKSGVLEHDMGWTALMACYKGRPDVY 120  
 QY 121 ELLSHGANSVYGLQYVYPIIWAAGRHADIVHLLONGAKVNCSDKGTGTPLVMAAR 180  
 DB 121 ELLSHGANSVYGLQYVYPIIWAAGRHADIVHLLONGAKVNCSDKGTGTPLVMAAR 180  
 QY 181 KGHLECYKHLLAGAVDQDQANSMTALYAVKGGTQSVKELKKNPNVNLTDKQGNRA 240  
 DB 181 KGHLECYKHLLAGAVDQDQANSMTALYAVKGGTQSVKELKKNPNVNLTDKQGNRA 240  
 QY 241 LMAESKEGHEIYQDLDAGTYVNIPIRSGDEVLIAVAGGHEIYRALLQKADIDIRG 300  
 DB 241 LMAESKEGHEIYQDLDAGTYVNIPIRSGDEVLIAVAGGHEIYRALLQKADIDIRG 300  
 QY 301 QDNKKTALYAVEKGNATWVDILQCNPDTEICTKDETPLIKATKKNRNIYVELLDKGA 360  
 DB 301 QDNKKTALYAVEKGNATWVDILQCNPDTEICTKDETPLIKATKKNRNIYVELLDKGA 360  
 QY 361 KVASAVDKKGDTPHVAIRGRSRRLAELLRNPKDGRILYRPNKAGETPINIDCSHOKSL 420  
 DB 361 KVASAVDKKGDTPHVAIRGRSRRLAELLRNPKDGRILYRPNKAGETPINIDCSHOKSL 420  
 QY 421 TQIFGARHLSPTEIDGMDLGYDLYSSALADILSEPTMOPICVGLYAQWGSCKSFLKKL 480  
 DB 421 TQIFGARHLSPTEIDGMDLGYDLYSSALADILSEPTMOPICVGLYAQWGSCKSFLKKL 480  
 QY 481 EDEMKTFAGQOTEPLOFQSWLIVFLLLCGGLGLVAPAPVDNTLAIASLSFLALYIF 540  
 DB 481 EDEMKTFAGQOTEPLOFQSWLIVFLLLCGGLGLVAPAPVDNTLAIASLSFLALYIF 540  
 QY 541 FIYIYRGGRGEGSMNMAWALSTRLARHIGYLELKLMEVNPPELPEOTTKLAPRPIF 600  
 DB 541 FIYIYRGGRGEGSMNMAWALSTRLARHIGYLELKLMEVNPPELPEOTTKLAPRPIF 600  
 QY 601 TDYNRULSVGGETSLEAMITATLSDACEREFGLATRLFRVFTRESOGKKMKTKCLPS 660  
 DB 601 TDYNRULSVGGETSLEAMITATLSDACEREFGLATRLFRVFTRESOGKKMKTKCLPS 660  
 QY 661 FVIFLEFIVGIIINGITLAEFRVDPKHLVATNLITIASVYGLAFVINCNTMMQVLDL 720  
 DB 661 FVIFLEFIVGIIINGITLAEFRVDPKHLVATNLITIASVYGLAFVINCNTMMQVLDL 720  
 QY 721 NSQRKRLHSAASKLHKLSEGFVKLCEVELMARMAKTIIDSTONOTRLVLIIDGLDAC 780  
 DB 721 NSQRKRLHSAASKLHKLSEGFVKLCEVELMARMAKTIIDSTONOTRLVLIIDGLDAC 780  
 QY 781 EODKVLQMDLTVRVLSEKGFPAIFASDPHIIITAIKAIQNLNSVLRDSNINCHDMRYIYH 840  
 DB 781 EODKVLQMDLTVRVLSEKGFPAIFASDPHIIITAIKAIQNLNSVLRDSNINCHDMRYIYH 840  
 QY 841 LPVFLNRSGLSNARKPLVTSATNGDITCSPTTGOEDTDYRONSIGEMTKIGSTALN 900  
 DB 841 LPVFLNRSGLSNARKPLVTSATNGDITCSPTTGOEDTDYRONSIGEMTKIGSTALN 900  
 QY 901 RRDYRRRQOQRTITROMSFDLTKLVTEDMFSDISPOTRRLINITYSVGRLLRANQIT 960  
 DB 901 RRDYRRRQOQRTITROMSFDLTKLVTEDMFSDISPOTRRLINITYSVGRLLRANQIT 960  
 QY 961 FNMORLASWINLTEOWPRTSWLILVLEETEGLPDQMTLTKMYERISKNIPTTKDVEPLL 1020  
 DB 961 FNMORLASWINLTEOWPRTSWLILVLEETEGLPDQMTLTKMYERISKNIPTTKDVEPLL 1020  
 QY 1021 EINGDINFEVFLSSRTPVLVARDVKTFLPCYVNLDPKLEIITADVRAAEQINIGLAV 1080  
 DB 1021 EINGDINFEVFLSSRTPVLVARDVKTFLPCYVNLDPKLEIITADVRAAEQINIGLAV 1080  
 QY 1081 PPLPLHAGPPRPSPGYQOPASVCSASFNGPPGVVSPPHSSYVSGLSGPPHPEFNRA 1140  
 DB 1081 PPLPLHAGPPRPSPGYQOPASVCSASFNGPPGVVSPPHSSYVSGLSGPPHPEFNRA 1140  
 QY 1141 AVPATGSSLLSSMTVDVCEKLRQIEGLDQNMMPQYCTTIKANINGRVLSCQNDIDELK 1200

DB 1141 SGAPAGPVVNLNSLVNDVAVCKEIKQIEGLDSMLPQYCTTIKANINGRVLSCQNDIDELK 1200  
 QY 1201 KEMANFCDWHLFNSMVLKMSVSVQVPEDPRLNENSAAPVHGESARSSSTELPLT 1260  
 DB 1201 KEMANFCDWHLFNSMVLKMSVSVQVPEDPRLNENSAAPVHGESARSSSTELPLT 1260  
 QY 1261 ELSQTPYTLNFSPELNTLIGDEGAPRHSNLSMOSQTRRTPSLSLKSDSSIEISKL 1320  
 DB 1261 ELSQTPYTLNFSPELNTLIGDEGAPRHSNLSMOSQTRRTPSLSLKSDSSIEISKL 1320  
 QY 1321 DKVOAEYRDAYREYIAQMSOLEGGTSGSTISGRSSPHSTVYIGOSSGSIHSTLEOERG 1380  
 DB 1321 DKVOAEYRDAYREYIAQMSOLEGGTSGSTISGRSSPHSTVYIGOSSGSIHSTLEOERG 1380  
 QY 1381 KEGELKQEDGKRSFLMKRGVDVYSSSGVSTNEMASPLDPIEPEKSPQSGSKLLPGKKS 1440  
 DB 1381 KEGELKQEDGKRSFLMKRGVDVYSSSGVSTNEMASPLDPIEPEKSPQSGSKLLPGKKS 1440  
 QY 1441 SERPSLEFQDCLKKGGGLRYOKLPDEDESGTGRVQITPHCSKMITRRLKAKORECASP 1500  
 DB 1441 SERPSLEFQDCLKKGGGLRYOKLPDEDESGTGRVQITPHCSKMITRRLKAKORECASP 1500  
 QY 1501 QEHSAPERTFTKAEVLSDALLDKSSDSGVRNNESSPNHSLHNEAADDSQLEKANLI 1560  
 DB 1501 QEHSAPERTFTKAEVLSDALLDKSSDSGVRNNESSPNHSLHNEAADDSQLEKANLI 1560  
 QY 1560 ELEDGHSKRCRMPHSISGLDPIIARMSIGSEDMKSPSECSLASSPEEMPCOKRYN 1619  
 DB 1560 ELEDGHSKRCRMPHSISGLDPIIARMSIGSEDMKSPSECSLASSPEEMPCOKRYN 1619  
 QY 1621 LNRTSTVTLNNNTAPTRANQNFDEIGIRETSQVILRPGPSNPPTAVQENLKSMAHK 1680  
 DB 1621 LNRTSTVTLNNNTAPTRANQNFDEIGIRETSQVILRPGPSNPPTAVQENLKSMAHK 1680  
 QY 1681 RSQRSSYTRLKSDASELH-AASSESTGGEEREESTL 1715  
 DB 1681 RSQRSSYTRLKSDASELH-AASSESTGGEEREESTL 1715  
 RESULT 4  
 AAM39025  
 ID AAM39025 standard; Protein; 1715 AA.  
 AC AAM39025;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 2170.  
 XX  
 KW Human; noctropic; immunosuppressant; cytosstatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemolactic;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000MO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.











Db 387 QIFGARHLSPETDGMVLGYDLYSSALADILSEPTMOPICVLYAOMGSGSKFLKLE 446  
QY 482 DEKKTTRAGQOTEPFLQPSWLIIVFLTLGGLGLVAFVPTDNLAIISLFLALYIF 541  
Db 447 DEKKTTRAGQOTEPFLQPSWLIIVFLTLGGLGLVAFVPTDNLAIISLFLALYIF 506  
QY 542 IVYFGRGRGESMNNAMALSTRARHIGYELLFKIMVNPPELPEOTTKALPVAFLE 601  
Db 507 IVYFGRGRGESMNNAMALSTRARHIGYELLFKIMVNPPELPEOTTKALPVAFLE 566  
QY 602 DYNRLSSVGETSLAEMIALSDACEREGFLATRLFRVFRTEESOGKKRKKKT 655  
Db 567 DYNRLSSVGETSLAEMIALSDACEREGFLATRLFRVFRTEESOGKKRKKKT 620  
RESULT 7  
ID AAE01035 standard; Protein: 551 AA.  
AAE01035:  
04-JUL-2001 (first entry)  
Human death domain-containing receptor (DDCR) protein from HODDX59 clone.  
Human: death domain-containing receptor; DDCR; immunosuppressive;  
antiarthritic; antirheumatic; antiproliferative; cytostatic;  
cardiac; vasotropic; cerebroprotective; neurotropic; neuroprotective;  
antibacterial; virucide; fungicide; ophthalmological; gene therapy;  
immunodeficiency disease; Acquired immune deficiency syndrome;  
AIDS; leukaemia; autoimmune disease; systemic lupus erythematosus;  
hyperproliferative disorder; neoplasm; cerebrovascular disorder;  
cerebral ischaemia; angiogenesis; cardiovascular disorder;  
neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
ocular disorder; corneal infection; degenerative disease; SMA; apoptosis;  
spinal muscular atrophy; epithelial cell proliferation; infection;  
cancer; wound healing; skin aging; chemotaxis; HODDX59 clone.  
OS Homo sapiens.  
XX  
FH Key location/Qualifiers  
FT Domain 399..415  
FT /label=Transmembrane\_domain  
FT /label=426..442  
FT Domain 64..69  
FT /label=Transmembrane\_domain  
FT /label=131..136  
FT Region 164..169  
FT /label=Immunogenic\_epitope  
FT Region 197..203  
FT /label=Immunogenic\_epitope  
FT Region 229..235  
FT /label=Immunogenic\_epitope  
FT Region 298..316  
FT /label=Immunogenic\_epitope  
FT Region 328..335  
FT /label=Immunogenic\_epitope  
FT Region 445..452  
FT /label=Immunogenic\_epitope  
FT Region 482..490  
FT /label=Immunogenic\_epitope  
FT Region 540..551  
FT /label=Immunogenic\_epitope  
XX  
PN WO200129063-A2.  
XX  
PD 26-APR-2001.  
XX  
PF 17-OCT-2000; 2000WO-US28666.  
XX  
PR 18-OCT-1999; 99US-0159585.

PR 24-NOV-1999; 99US-0167246.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI N1 J, Ruben SM;  
XX  
DR WPI: 2001-300316/31.  
XX  
N-PSDB: AAD04776.  
PT New death domain containing receptor polynucleotides and polypeptides,  
useful for treating and diagnosing cancer.  
PS Claim 1; Page 292-294; 298pp; English.  
XX  
CC The patent discloses novel death domain-containing receptor (DDCR)  
CC cDNAs and their corresponding proteins. DDCR cDNA and protein are  
CC used to prevent, treat or ameliorate a medical condition in mammals.  
CC They are also used in diagnosing a pathological condition or  
CC susceptibility to a pathological condition. The DDCR protein and  
CC its antibodies are used in the diagnosis and treatment of disorders  
CC such as immunodeficiency diseases (e.g. Acquired immune deficiency  
CC syndrome (AIDS), leukaemia) autoimmune diseases (e.g. systemic lupus  
CC erythematosus, rheumatoid arthritis), hyperproliferative disorders  
CC (e.g. neoplasms of the breast or liver), cerebrovascular disorders  
CC (e.g. cerebral ischaemia, angiogenesis), cardiovascular disorders  
CC (e.g. cardiac arrest), neurodegenerative diseases (e.g. Alzheimer's  
CC disease, Parkinson's disease), ocular disorders (e.g. corneal  
CC infection), degenerative diseases (e.g. spinal muscular atrophy-SMA),  
CC cancer, aberrant apoptosis, disorders of the placenta or uterus and  
CC infections caused by bacteria, viruses and fungi. The DDCR proteins  
CC are used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, to support cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The DDCR sequences are used  
CC in correcting aberrant cellular apoptosis by gene therapy.  
CC The present sequence is human death domain-containing receptor  
CC (DDCR) protein from clone HODDX59.  
XX  
SQ Sequence 551 AA:  
Query Match 31.4%; Score 2782.5; DB 22; Length 551;  
Best Local Similarity 96.7%; Pred. No. 3e-210; 9; Indels 1; Gaps 1;  
Matches 534; Conservative 8; Mismatches 9;  
QY 102 MGSWTALMAGCYKGRDVTVELLSHGANSVGLGYVPIIIMAGRGHADVHLLONG 161  
Db 1 MGSWTALMAGCYKGRDVTVELLSHGANSVGLGYVPIIIMAGRGHADVHLLONG 59  
QY 162 AKVNSDKYGTPLVMAARKGHLECYKHLLAMGADVDGANSMTALIVAVGGYTOSYK 221  
Db 60 AKVNSDKYGTPLVMAARKGHLECYKHLLAMGADVDGANSMTALIVAVGGYTOSYK 119  
QY 222 ELIKRPNVNLTKDQNTALMTASKGHEIYODLIDATYVNPDRSGDYLICAVRG 281  
Db 120 ELIKRPNVNLTKDQNTALMTASKGHEIYODLIDATYVNPDRSGDYLICAVRG 179  
QY 282 HVEIVALLQKYADIDIRGQDNKTALYMAVENGNATVBDILQCNPDTEICTKDETPLI 341  
Db 180 HVEIVALLQKYADIDIRGQDNKTALYMAVENGNATVBDILQCNPDTEICTKDETPLI 239  
QY 342 KATKMRNIEVELLDKGAQVSAVDKGTPLHVAIRGSRRLAELLRNPDGRILYRP 401  
Db 240 KATKMRNIEVELLDKGAQVSAVDKGTPLHVAIRGSRRLAELLRNPDGRILYRP 299  
QY 402 NKAGETPVNIDCSHOKSILTOIFGARHLSPETDGMVLGYDLYSSALADILSEPTMOPPI 461  
Db 300 NKAGETPVNIDCSHOKSILTOIFGARHLSPETDGMVLGYDLYSSALADILSEPTMOPPI 359  
QY 462 CVGLYAOMGSGSKFLKLEDEKKTTRAGQOTEPFLQPSWLIIVFLTLGGLGLVAFV 521  
Db 360 CVGLYAOMGSGSKFLKLEDEKKTTRAGQOTEPFLQPSWLIIVFLTLGGLGLVAFV 419  
QY 522 DTNLAIINISLFLALYIFFIYIVYFGRGRGESMNNAMALSTRARHIGYELLFKIMV 581

DB 420 HNNLGIATVSLSEFALTYIFIVYFGREGESMMWAWLSTRLARHIGYLELLKMEV 479  
 OY 582 NPPELPEOTTKALPVAFLEFTDYNRLSSVGGESTSLAEMTATLSDAEEREGFATLFRVF 641  
 DB 480 NPPELPEOTTKALPVAFLEFTDYNRLSSVGGESTSLAEMTATLSDAEEREGFATLFRVF 539  
 OY 642 RTEESGKKKKWK 653  
 DB 540 KTEDTGKKKKK 551

RESULT 8  
 AAB95191  
 ID AAB95191 standard; Protein; 543 AA.  
 AC AAB95191;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:17273.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 XX  
 PR 27-AUG-1999; 99JP-0300253.  
 XX  
 PR 11-JAN-2000; 2000JP-0118776.  
 XX  
 PR 02-MAY-2000; 2000JP-0183767.  
 XX  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 Claim 8; SEQ ID 17273; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide which comprises a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.  
 XX  
 SQ Sequence 543 AA;  
 Query Match 27.2%; Score 2408; DB 22; Length 543;  
 Best Local Similarity 86.8%; Pred. No. 1e-180;  
 Matches 472; Conservative 24; Mismatches 46; Indels 2; Gaps 2;

OY 1173 MNPYCTTICKANINGVLSQCNIDELKKEAMNFGDMHFRSMVLEMSVEQVYPPDP 1232  
 DB 1 MLPQYCTTICKANINGVLSQCNIDELKKEAMNFGDMHFRSMVLEMSVEQVYPPDP 60  
 OY 1233 RFLNENSSAPVPHEGSEARRSSHTLPTLELSQPTPLNFSFELNTLIGDEAPRHSNL 1292  
 DB 61 RFLSESSSGPAPHEGSEARRASHNLPTLELSQPTPLNFSFELNTLIGDEAPRHSNL 120  
 OY 1293 SWSQTRRTPLSLNSQDSSEISLTKTDVQAEYRDARYEYIAQMSQLEGSGSTTIG 1352  
 DB 121 SWSQTRRTPLSLNSQDSSEISLTKTDVQAEYRDARYEYIAQMSQLEGSGSTTIG 180  
 OY 1353 RSPHSTYYIGOSSSGSISHTLEOERKGEELKOEGRKSFILMKRGDVIYSSGVSIN 1412  
 DB 181 RSPHSTYYIGOSSSGSISHTLEOERKGEELKOEGRKSFILMKRGDVIYSSGVSIN 240  
 OY 1413 EASPLDPTFEDEKSDSGSKLLPGKKSSEPSLFTQDLKLGGLRYOKLPSEDESGT 1472  
 DB 241 DASPLDPTFEDEKSDSGSKLLPGKKSSEPSLFTQDLKLGGLRYOKLPSEDESGT 300  
 OY 1473 GRVQITPHCSKMTKTKAKKORECASPOHSAEPIRTFKAKETYSDALDKKSSDSG 1532  
 DB 301 EESDNTF-LTKDKDRKAEKVERVPSPEHSAEPIRTFKAKETYSDALDKKSSDSG 359  
 OY 1533 VRSNESPNSLNEAADDSOLKANKLIELEDGSHGKGMPSLSGLDPTIARRSTCS 1592  
 DB 360 VRSSESPNSLNEAADDSOLKANKLIELEDGSHGKGMPSLSGLDPTIARRSTCS 419  
 OY 1593 EDKSPSECSLIASSPEESWPAQOKAYNLRRPSTYTLNNTAPTRANQNEDELIGRE 1652  
 DB 420 EDKSPSECSLIASSPEESWPAQOKAYNLRRPSTYTLNNTAPTRANQNEDELIGRE 479  
 OY 1653 TSOVYILRPGSPWPVAVONNLKSMARKSQRSSYTRLSKASELH-AASSESTGGEER 1711  
 DB 480 TSOVYILRPGSPWPVAVONNLKSMARKSQRSSYTRLSKASELH-AASSESTGGEER 539  
 OY 1712 ESIL 1715  
 DB 540 ESIL 543

RESULT 9  
 ABB64857  
 ID ABB64857 standard; Protein; 1498 AA.  
 AC ABB64857;  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 21363.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 PN WO200171042-A2.  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX



KM inflammatory disorder; acquired immune deficiency syndrome; AIDS;  
 KM Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;  
 KM Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;  
 KM glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;  
 KM osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;  
 KM infection; cell proliferative disorder; actinic keratosis; myeloma;  
 KM arteriosclerosis; notropic; anticonvulsant; antithyroid; nephrotropic;  
 KM neuroprotective; dermatological.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 94..108 /note="Adrenocorticotrophin receptor"  
 FT Domain 196..198 /note="Rgd cell interaction motif"  
 FT Region 328..335 /note="Phage integrase family"  
 FT  
 WO200142285-A2.  
 XX  
 14-JUN-2001.  
 XX  
 05-DEC-2000; 2000WO-US32990.  
 PF  
 10-DEC-1999; 99US-0172852.  
 PR 16-DEC-1999; 99US-0172354.  
 XX  
 (INCY-) INCYTE GENOMICS INC.  
 PI Yue H, Tang YT, Lal P, Burford N, Azimzal Y, Patterson C;  
 PI Baughn MR, Lu DM, Shah P, Au-Young J;  
 XX  
 DR MPI: 2001-381632/40.  
 DR N-PSDB; AAD08053.  
 XX  
 New human extracellular matrix and cell adhesion molecules and  
 PT polynucleotide sequences encoding them, useful for diagnosis,  
 PT prevention, treatment of genetic, autoimmune and cell proliferative  
 PT disorders.  
 XX  
 PS Claim 1; Page 101-102; 135pp; English.  
 XX  
 The present sequence is a human extracellular matrix and cell  
 CC adhesion molecule (XMAP). The XMAP is used for screening a compound for  
 CC effectiveness as an agonist or antagonist of XMAP. The identified agonist  
 CC or antagonist are used for treating a disease or condition associated  
 CC with decreased or increased expression of functional XMAP. The  
 CC polynucleotides encoding XMAP are useful in somatic or germ-line gene  
 CC therapy to correct a genetic deficiency, to express a conditionally  
 CC lethal gene product and to express a protein which affords protection  
 CC against intracellular parasites and also for diagnosis of disorders  
 CC associated with expression of XMAP. They are also used for generating  
 CC hybridisation probes useful in mapping the naturally occurring genomic  
 CC sequences and to create knock in humanised animals (pigs) or transgenic  
 CC animals (mice or rats) to model human diseases. Oligonucleotide or longer  
 CC fragments derived from the polynucleotide sequences may be used as  
 CC elements on a microarray. Antibodies which specifically bind XMAP may be  
 CC used for the diagnosis of disorders associated with the expression of  
 CC XMAP, or in assays to monitor patients being treated with XMAP. Diseases  
 CC diagnosed, prevented or treated include genetic disorders such as  
 CC adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's  
 CC disease, myotonic dystrophy, sickle cell anaemia, thalassemia,  
 CC autoimmune/inflammatory disorders such as acquired immune deficiency  
 CC syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,  
 CC atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,  
 CC glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,  
 CC osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,  
 CC bacterial, fungal, parasitic, protozoal and helminthic infections and  
 CC cell proliferative disorders such as actinic keratosis, arteriosclerosis  
 CC and cancer including breast, bladder, bone marrow, brain and uterus  
 CC cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.  
 XX

SO Sequence 513 AA;  
 Query Match 25.5%; Score 2260; DB 22; Length 513;  
 Best Local Similarity 86.6%; Pred. No. 4.4e-169;  
 Matches 445; Conservative 21; Mismatches 46; Indels 2; Gaps 2;  
 QY 1203 MANNFSDWHLFRSMVLEMSVESQVYDEPREFLENSSAPVPHGESARRSHTLEL 1262  
 DB 1 MNNFQDWHLFRSTVLEMAHSHVYEDPREFLESSESSGAPVHGEARRASNNELPHTEL 60  
 QY 1263 SSOTPYTLNFSFELMTLGLDEGAPRHSNLSWMSQTRRPSSLNSQSSSTEISKLDK 1322  
 DB 61 SSOTPYTLNFSFELMTLGLDEGAPRHSNLSWMSQTRRPSSLNSQSSSTEISKLDK 120  
 QY 1323 VOAEYRDAYREYIAQMSOLEGCGSTTIGRSSPHSTTYIGQSSGSGSTHSTLEDEGRKE 1382  
 DB 121 VOAEYRDAYREYIAQMSOLEGCGSTTIGRSSPHSTTYIGQSSGSGSTHSTLEDEGRKE 180  
 QY 1383 GELKQEDGRKSFLLMKRGVDYIDYSSGVSSTNEASPLDPIEDEKSDQSGSKLLPKKSS 1442  
 DB 181 SEKPPDGRKSFLLMKRGVDYIDYSSGVSSTNEASPLDPIEDEKSDQSGSKLLPKKSS 240  
 QY 1443 RPSLFQTDLKLKGGGLRYOKLPSDEDESGTGRVQITPHCSKMITRKLAKQRECAPOE 1502  
 DB 241 RPSLFQTDLKLKGGGLRYOKLPSDEDESGTGRVQITPHCSKMITRKLAKQRECAPOE 299  
 QY 1503 HSAEPIRTFKAKEYLSDALLDKKSDSGVRSNESSPMHSLHNEAADSOLEKANLTEL 1562  
 DB 300 HSAEPIRTFKAKEYLSDALLDKKSDSGVRSNESSPMHSLHNEAADSOLEKANLTEL 359  
 QY 1563 EDEGHSGRKMPHSLSLGLDPIIARMSISEDKKSPSECSLIASSPEESWPAQORAYNLN 1622  
 DB 360 EDDSHSGKRGIRPHSLSLGLDPIIARMSISEDKKSPSECSLIASSPEESWPAQORAYNLN 419  
 QY 1623 RTPSTYTLNNTAPTRANQNPDEIGIETSYVILRPQSPNPTRAVQENELKNAHKRS 1682  
 DB 420 RTPSTYTLNNTAPTRANQNPDEIGIETSYVILRPQSPNPTRAVQENELKNAHKRS 479  
 QY 1683 QRSSYTRLSKDASELH-AASSESTGFEERESTL 1715  
 DB 480 QRSSYTRLSKDPPELHMAASSESTGFEERESTL 513  
 XX  
 RESULT 11  
 ABG08697  
 ID ABG08697 standard; Protein; 624 AA.  
 XX  
 AC ABG08697;  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #8688.  
 XX  
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR MPI: 2001-639362/73.  
 DR N-PSDB; AAS72884.  
 XX





Db 1 TOSVKEILKRNPNVNLTDKGNLTALMISKEGHTEIVODLDAGTYVNPDRSGTYLIG 60  
QY 277 AARGHVEIVRALLQKYADIDIRGDNKNTALYMAVEKGNATWRILOCNPDTEICTYDG 336  
61 AARGHVEIVRALLQKYADIDIRGDNKNTALYMAVEKGNATWRILOCNPDTEICTYDG 120

## RESULT 13

AAB94785  
ID AAB94785 standard; Protein; 129 AA.

AC AAB94785;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:15890.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

XX EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs.

PS Claim 8; SEQ ID 15890; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
XX full-length cDNAs defined in the specification. Where a primer set  
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
XX to the complementary strand of a polynucleotide which comprises one of  
XX the 5602 nucleotide sequences defined in the specification, where the  
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
XX of an oligonucleotide comprising a sequence complementary to the  
XX complementary strand of a polynucleotide which comprises a 5'-end  
XX sequence and an oligonucleotide comprising a sequence complementary to a  
XX polynucleotide which comprises a 3'-end sequence, where the  
XX oligonucleotide comprises at least 15 nucleotides and the combination  
XX of the 5'-end sequence/3'-end sequence is selected from those defined in  
XX the specification. The primer sets can be used in antisense therapy and  
XX in gene therapy. The primers are useful for synthesizing polynucleotides,  
XX particularly full-length cDNAs. The primers are also useful for the  
XX detection and/or diagnosis of the abnormality of the proteins encoded by  
XX the full-length cDNAs. The primers allow obtaining of the full-length  
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
XX represent oligonucleotides, all of which are used in the exemplification  
XX of the present invention.

XX Sequence 129 AA;

XX Query Match

6.8%; Score 599.5; DB 22; Length 129;

Best Local Similarity 90.7%; Pred. No. 4.2e-39;  
Matches 117; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1588 MSICSEDKKSPSECLSIASSPESWPAQKAYNLNRTPTVTNNNTAPNRANQFDEI 1647  
Db 1 MSICSEDKKSPSECLSIASSPESWPAQKAYNLNRTPTVTNNNTAPNRANQFDEI 60

QY 1648 EGIRETSQVILRPPSPNPTAPVONENIKSMAHRSQRSSYTRISKASELH-AASESG 1706  
Db 61 EGIRETSQVILRPPSPNPTAPVONENIKSMAHRSQRSSYTRISKASELH-AASESG 120

QY 1707 FGEERESIL 1715  
Db 121 FGEERESIL 129

## RESULT 14

AAB56277  
ID AAB56277 standard; Protein; 109 AA.

AC AAB56277;

DT 13-MAR-2001 (first entry)

DE Human secreted protein sequence encoded by gene 82 SEQ ID NO:371.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
XX antiinflammatory; antiproliferative; cytostatic; cardiant; vasotropic;  
XX cerebroprotective; neurotropic; neuroprotective; antibacterial; vitucide;  
XX fungicide; ophthalmological; gene therapy; pathological condition;  
XX autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;  
XX neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischemia;  
XX cerebrovascular disorder; angiogenesis; nervous system disorder;  
XX Alzheimer's disease; infection; ocular disorder; corneal infection;  
XX wound healing; skin aging; food additive; preservative.

OS Homo sapiens.

XX WO200070042-A1.

XX 23-NOV-2000.

XX 11-MAY-2000; 2000WO-US12788.

XX 13-MAY-1999; 99US-0134068.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;  
PI Duan RD, Florence KA, Soppet DR;  
DR WPI; 2000-679828/66.

PT Isolated nucleic acid molecule encoding a human secreted protein is  
PT used in preventing, treating or ameliorating a medical condition -  
PS Disclosure; Page 1025-1026; 1065bp; English.

XX The polynucleotide sequences given in AAC99818 to AAC99977 encode the  
XX human secreted proteins given in AAB56077 to AAB56362. Human secreted  
XX proteins have activities based on the tissues and cells the genes are  
XX expressed in. Examples of activities include: immunosuppressive;  
XX antiarthritic; antiinflammatory; antiproliferative; cytostatic; cardiant;  
XX vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;  
XX vitucide; fungicide; and ophthalmological. The human secreted  
XX polynucleotides and proteins can be used to prevent, treat or ameliorate  
XX a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,  
XX dogs, chickens or sheep. They are also used in diagnosing a pathological  
XX condition or susceptibility to a pathological condition. Disorders which  
XX are diagnosed or treated include autoimmune diseases e.g. rheumatoid  
XX arthritis, hyperproliferative disorders e.g. neoplasms of the breast or  
XX liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular  
XX disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders



to determine if any of the above promoters can promote the expression and/or function of the gene of interest.

[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 55.5 Seconds

(without alignments)  
4117.564 Million cell updates/sec

Title: US-10-021-571-4

Perfect score: 8884

Sequence: 1 MSVLISQSVINVEENIPD.....LHAASSSTGGERESIL 1715

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8884	100.0	1715	22	AAAM38993
2	8884	100.0	1715	22	AAAM39025
3	8884	100.0	1715	23	AAU96840
4	8362.5	94.1	1762	23	AAU96841
5	8301	93.4	1763	23	AAU80244
6	3050.5	34.3	705	22	AAU75604
7	2851.5	32.1	551	22	AAU75604
8	2791	31.4	543	22	AAU95191
9	2638	29.7	513	22	AAU95191
10	2503	28.2	624	22	ABG08697

11	2379.5	26.8	1498	22	ABB64857
12	666	7.5	129	22	AAU94785
13	620	7.0	120	22	AAU0811
14	553	6.2	109	21	AAU56277
15	526	5.9	1872	22	AAU79160
16	521	5.9	4274	22	ABG00972
17	519.5	5.8	2443	22	ABB60521
18	519.5	5.8	4386	22	ABG07375
19	519	5.8	1724	23	AAU020513
20	514	5.8	4397	22	ABG21944
21	499.5	5.6	187	22	AAU40779
22	482.5	5.4	784	22	ABB62206
23	480.5	5.4	1745	19	AAU70608
24	480.5	5.4	1745	19	AAU76776
25	480.5	5.4	1745	21	AAU11589
26	480.5	5.4	1745	23	AAU021368
27	478.5	5.4	1549	22	ABB58328
28	478.5	5.4	1549	22	ABB67412
29	463	5.2	2119	22	ABB64823
30	447	5.0	743	23	ABB97453
31	446	5.0	1133	22	ABG08071
32	440	5.0	378	22	ABB64963
33	438.5	4.9	1053	22	AAU28174
34	438	4.9	1088	22	ABB62429
35	438	4.9	1097	22	ABG21941
36	437	4.9	740	22	AAU01030
37	437	4.9	1569	23	AAU74360
38	430.5	4.8	705	22	AAU20496
39	429	4.8	435	22	AAU66710
40	429	4.8	435	22	AAU93879
41	429	4.8	627	23	AAU17136
42	421	4.7	1377	22	ABG08072
43	417.5	4.7	2441	22	ABB62231
44	415.5	4.7	1267	22	ABG22790
45	415	4.7	378	22	ABB64962

## ALIGNMENTS

RESULT 1	AAU38993	standard; Protein; 1715 AA.
ID	AAU38993	
XX	AAU38993	
AC	AAU38993	
XX	22-OCT-2001	(first entry)
DT	22-OCT-2001	
DE	Human polypeptide SFG ID NO 2138.	
XX	Human; nocitropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.	
OS	Homo sapiens.	
XX	WO200153312-A1.	
PN	26-JUL-2001.	
PD	26-JUL-2001.	
XX	26-DEC-2000; 2000MO-US34263.	
PF	21-JAN-2000; 2000US-0488725.	
XX	25-APR-2000; 2000US-0552317.	
PR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	



QY 1621 NRPESTVTLNNSAPANRANQNFDEMEGIRETSQVILRPSSPNPTTIONENLKSMTKRR 1680  
 DB 1621 NRPESTVTLNNSAPANRANQNFDEMEGIRETSQVILRPSSPNPTTIONENLKSMTKRR 1680  
 QY 1681 SQRSSTYRLSKDPELHAASSESTGGEERESTL 1715  
 DB 1681 SQRSSTYRLSKDPELHAASSESTGGEERESTL 1715

RESULT 2  
 ID AAM39025  
 ID AAM39025 standard; Protein; 1715 AA.  
 AC AAM39025;  
 XX 22-OCT-2001 (first entry)  
 DT Human polypeptide SEQ ID NO 2170.  
 DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 XX Peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KW Leukaemia.  
 XX Homo sapiens.  
 OS WO200153312-A1.  
 PN 26-JUL-2001.  
 PD 26-DEC-2000; 2000WO-US34263.  
 PE 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX (HYSE-) HYSEQ INC.  
 FA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPI: 2001-442253/47.  
 DR N-PSDB; AAI58181.  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 4; SEQ ID NO 2170; 10078pp; English.  
 XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AA42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 1715 AA:  
 SQ Query Match 100.0%; Score 8884; DB 22; Length 1715;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYLISQSVINVEEENIRALAKLEKCDVDERNCCGTPIMIAEGCNLEIVELIKN 60  
 DB 1 MSYLISQSVINVEEENIRALAKLEKCDVDERNCCGTPIMIAEGCNLEIVELIKN 60  
 QY GANCNLEDDNNMTALISASKEGHHVIEELKCGVNLERDNGMTALMACYKGRTPVY 120  
 DB GANCNLEDDNNMTALISASKEGHHVIEELKCGVNLERDNGMTALMACYKGRTPVY 120  
 QY 121 ELLISHGAMPSTGLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTPVWAAR 180  
 DB 121 ELLISHGAMPSTGLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTPVWAAR 180  
 QY 181 KGHLECVKHLAMGADVDGANSMTALVAVKGGTOSVKRILKRNPNVNLTDKDNAT 240  
 DB 181 KGHLECVKHLAMGADVDGANSMTALVAVKGGTOSVKRILKRNPNVNLTDKDNAT 240  
 QY 241 LMTASKEGHTIEYQDLDAGTYVNIIPDRSGDVIIGAVNGHVEIVRALLQKADIDIRG 300  
 DB 241 LMTASKEGHTIEYQDLDAGTYVNIIPDRSGDVIIGAVNGHVEIVRALLQKADIDIRG 300  
 QY 301 QDNKTALVYAVEKGNATVYRDILQCNPDTEICTKGETPLIKATKRNIEVELLDKGA 360  
 DB 301 QDNKTALVYAVEKGNATVYRDILQCNPDTEICTKGETPLIKATKRNIEVELLDKGA 360  
 QY 361 KYSADVKKGDDPLHAIIRGRSRKLAELLRNKDGLYRPAKAGETPNIDCSHOKSL 420  
 DB 361 KYSADVKKGDDPLHAIIRGRSRKLAELLRNKDGLYRPAKAGETPNIDCSHOKSL 420  
 QY 421 TQIFGARHLSPETDMDLGYDLYSSALADILSEPTMOPPICVGYAOWSGSKSFLKL 480  
 DB 421 TQIFGARHLSPETDMDLGYDLYSSALADILSEPTMOPPICVGYAOWSGSKSFLKL 480  
 QY 481 EDEMKTFACQOIEPLFOFSWLVLTLLCGGLGILPAFTVHPNIGIAVSFLALYIF 540  
 DB 481 EDEMKTFACQOIEPLFOFSWLVLTLLCGGLGILPAFTVHPNIGIAVSFLALYIF 540  
 QY 541 FIYIFGGRREGESENMAWVLSRLARHIGYELLLKMFVNPELPEDOTKALPVRELF 600  
 DB 541 FIYIFGGRREGESENMAWVLSRLARHIGYELLLKMFVNPELPEDOTKALPVRELF 600  
 QY 601 TDYNRLSVGGETSLAEMITATLSDACEREFGLATRLRFVFTEDTQKKKKKTCCLPS 660  
 DB 601 TDYNRLSVGGETSLAEMITATLSDACEREFGLATRLRFVFTEDTQKKKKKTCCLPS 660  
 QY 661 FVIFLEFICITISGITLLAIFRVDPKHLVNAVVISASVGLAVLNCRTMWOVLDSL 720  
 DB 661 FVIFLEFICITISGITLLAIFRVDPKHLVNAVVISASVGLAVLNCRTMWOVLDSL 720  
 QY 721 NSQKRHLNAAASKLHLKSEGEFMVKLCVEELAMAMAKTIDSFOTNORLVIYIDGLAC 780  
 DB 721 NSQKRHLNAAASKLHLKSEGEFMVKLCVEELAMAMAKTIDSFOTNORLVIYIDGLAC 780  
 QY 781 EDDKVLQMDIVRVLFSGKPTAIFASDPHIIKAIINONLSVLRDSNINHGYRNIVH 840  
 DB 781 EDDKVLQMDIVRVLFSGKPTAIFASDPHIIKAIINONLSVLRDSNINHGYRNIVH 840  
 QY 841 LVPFLNSRGLSNARKFLVTSANGVPCSDTGTGIEDDADRVRSQSLGEMKRLSGKTALN 900  
 DB 841 LVPFLNSRGLSNARKFLVTSANGVPCSDTGTGIEDDADRVRSQSLGEMKRLSGKTALN 900  
 QY 901 RRDYRRRQMGTTITROMSFDLTLLVTEWFSIDSPTMRRLNIVSVTRLLRANQIS 960  
 DB 901 RRDYRRRQMGTTITROMSFDLTLLVTEWFSIDSPTMRRLNIVSVTRLLRANQIS 960  
 QY 961 FHWDRLASINILTEOMPRTSWLLIYLBETBSIPQDMTKTYERISKINIPFYKQVEPL 1020  
 DB 961 FHWDRLASINILTEOMPRTSWLLIYLBETBSIPQDMTKTYERISKINIPFYKQVEPL 1020

```

Db      961  FNMRLASINILTEQWMPYRTSMILILEETBGIPIPDQMLTKTYERISKNIPTTKDVEPLL 1020
QY      1021  EIDGDINFEVFLSSRIPLVARYKVFLEPCTVNLDPKLEITADVAAAEQISIGLAY 1080
Db      1021  EIDBDIRNFEEVLSRIPLVARYKVFLEPCTVNLDPKLEITADVAAAEQISIGLAY 1080
QY      1081  PPLPLHEGPAPRAGSYGSPFVSCSTSFNGFAGGVVSPHSSYSGMTGPPQPFNRG 1140
Db      1081  PPLPLHEGPAPRAGSYGSPFVSCSTSFNGFAGGVVSPHSSYSGMTGPPQPFNRG 1140
QY      1141  SGAPGPPVLLNSLNDVAVCEKIKQIEGLDQSMLEPOYCTTIKANKINGRYLAQCNDLEK 1200
Db      1141  SGAPGPPVLLNSLNDVAVCEKIKQIEGLDQSMLEPOYCTTIKANKINGRYLAQCNDLEK 1200
QY      1201  KEMNNFGDWHLFSTYLEMNAASHVYPPEDPRLSSSSGAPAHGEPARASINELPHT 1260
Db      1201  KEMNNFGDWHLFSTYLEMNAASHVYPPEDPRLSSSSGAPAHGEPARASINELPHT 1260
QY      1261  ELSQTPYTLNFSFEELNTGLDGAAPRHSNLSMOSQTRTPSLSLNSODSIEISKLT 1320
Db      1261  ELSQTPYTLNFSFEELNTGLDGAAPRHSNLSMOSQTRTPSLSLNSODSIEISKLT 1320
QY      1321  DKVOAERYDAYREYIAQMSOLEGGPGSTTISGRSSPHSTYYMGQSSGGSIHNLDEK 1380
Db      1321  DKVOAERYDAYREYIAQMSOLEGGPGSTTISGRSSPHSTYYMGQSSGGSIHNLDEK 1380
QY      1381  KDSEKPPDGRKSPFLMKRGVYIDYSSGVSSTNDASPLDPTTEDEKSDQSGSKLLPKKKS 1440
Db      1381  KDSEKPPDGRKSPFLMKRGVYIDYSSGVSSTNDASPLDPTTEDEKSDQSGSKLLPKKKS 1440
QY      1441  SERSLFQTDKLKSGILRYOKLPSEDESGTEESDMTPLLKDKRKAKEGVYRVPKSP 1500
Db      1441  SERSLFQTDKLKSGILRYOKLPSEDESGTEESDMTPLLKDKRKAKEGVYRVPKSP 1500
QY      1501  EHSAPRTFTKAREYISDALLDKKSDSGVRSSESSPNLSLNEVADDSQLEKANLIE 1560
Db      1501  EHSAPRTFTKAREYISDALLDKKSDSGVRSSESSPNLSLNEVADDSQLEKANLIE 1560
QY      1561  LEDDSHGKRGIPHSLSGLDPIIARMSICSEDKKSPSECSLIASSPEENPACOKAYNL 1620
Db      1561  LEDDSHGKRGIPHSLSGLDPIIARMSICSEDKKSPSECSLIASSPEENPACOKAYNL 1620
QY      1621  NRTPTVTLLNNSAPANRANONFDEMEGIRETSOYILRPSSTPPTIIONENLKSMTIKR 1680
Db      1621  NRTPTVTLLNNSAPANRANONFDEMEGIRETSOYILRPSSTPPTIIONENLKSMTIKR 1680
QY      1681  SORSSYTRLSKDPPELHAASSESTGFEERESIL 1715
Db      1681  SORSSYTRLSKDPPELHAASSESTGFEERESIL 1715

RESULT 3
AAU96840 standard: Protein; 1715 AA.
AAU96840;
30-JUL-2002 (first entry)
Human kidins220 protein.
XX kidins220; kinase D interacting substrate of 22kDa; cytosolic;
XX neuroprotective; gene therapy; protein kinase D; PKD; cancer;
XX neurodegenerative disease; glioblastoma multiforme; prostate cancer;
XX human.
OS Homo sapiens.
XX
XX WO200220786-A2.
XX
XX 14-MAR-2002.
XX
XX 06-SEP-2001; 2001WO-GB03977.

```

```

XX      06-SEP-2000; 2000US-230449P.
PR      (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.
PA      Schiavo G, Iglesias T;
XX      WPI: 2002-371879/40.
XX      N-PSDB; ABR51221.
DR      Polypeptide kinase D interacting substrate of 220 kDa and
XX      polynucleotide, useful for identifying modulators useful in treating
XX      cancer and neurodegenerative diseases
PS      Claim 33; Fig 11; 228pp; English.
CC      The invention relates to polypeptide comprising a kinase D interacting
CC      substrate of 220 kDa (kidins220) from rat and human or their
CC      fragment, variant or fusion provided that the protein is not the
CC      polypeptide encoded by Genbank Accession No. AB033076. Also included are
CC      a polynucleotide encoding kidins220 provided that polynucleotide is
CC      not any one of the clones corresponding to the 61 Genbank Accession Nos.
CC      given in the specification; an expression/replacable vector comprising
CC      the polynucleotide; a recombinant host cell containing the polynucleotide
CC      or vector; an anti-kidins220 antibody (used in the preparation of
CC      kidins220); an agent (A1) which modulates activity of kidins220;
CC      protein kinase D (PKD) or interaction between PKD and kidins220;
CC      a mutant animal transgenic for kidins220; the use of an agent capable of
CC      detecting the expression of kidins220 gene in manufacture of a diagnostic
CC      reagent for diagnosing or prognostic cancer or for monitoring the
CC      progression of cancer in a patient; a compound comprising a moiety that
CC      selectively binds to kidins220 or its variant and another moiety; and a
CC      nucleic acid molecule encoding the compound. The antibody is useful for
CC      modulating the activity of kidins220, where the antibody prevents a
CC      region of kidins220 interacting from another protein, or reduces the
CC      ability of kidins220 to bind to adenosine triphosphate (ATP). The
CC      kidins220 polynucleotide is useful for identifying an agent which
CC      modulates the promoter activity of the polynucleotide, and kidins220 is
CC      useful for identifying an agent which modulates activity of kidins220,
CC      protein kinase D (PKD) and the interaction between PKD with kidins220.
CC      kidins220 or its polynucleotide is useful in medicine, e.g. using gene
CC      therapy or for diagnosing or monitoring progression of cancer in a
CC      patient. The agent is useful in manufacture of medicament, for use in
CC      treatment of neurodegenerative disease, and the agent which inhibits
CC      function of kidins220 gene or its product is useful for treating cancer
CC      which is glioblastoma multiforme or prostate cancer. The compound is
CC      useful for imaging cancer in an individual, and for diagnosing or
CC      prognostic, and also for treating an individual. The present
CC      sequence represents Human kidins220.
SQ      Sequence 1715 AA:
XX
XX      Query Match 100.0%; Score 8884; DB 23; Length 1715;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MSYLSIQSVINVEEENIPALKALKLEKCDVDERNECGOTPLMAEAGNLEIVKELKN 60
Db      1 MSYLSIQSVINVEEENIPALKALKLEKCDVDERNECGOTPLMAEAGNLEIVKELKN 60
QY      61 GANCNLEDDNWTALISASKEGHVTVVEELKCGVLEHRDGGWALMWACYKGTQDV 120
Db      61 GANCNLEDDNWTALISASKEGHVTVVEELKCGVLEHRDGGWALMWACYKGTQDV 120
QY      121 ELLSHGANSYTGLOYSYPIIWAAGRHADIVHLLDNGAKVNCSDKTYGTPPLWMAAR 180
Db      121 ELLSHGANSYTGLOYSYPIIWAAGRHADIVHLLDNGAKVNCSDKTYGTPPLWMAAR 180
QY      181 KGHLECYKHLANGADVDEGANSKATLIVAVKGYTQSVKELKRNPNVNLTDKDGNTA 240
Db      181 KGHLECYKHLANGADVDEGANSKATLIVAVKGYTQSVKELKRNPNVNLTDKDGNTA 240
QY      241 LMTASKEGHEIYQDLDAGTYVNIIPDRSGDVTYLGAVAGHVEIYRALDQYADIDIRG 300

```



```

Db 241 LMIASKEGHEIYQDLDAGTYVNIIPDRSGDFVLICAVRGHVEIYRALLQYADIDING 300
Qy 301 QDNKTLAYAVENGNATWTRDILQCPDDEICTKGEFTPLKATKRNIEVEVLLDKGA 360
Db 301 QDNKTLAYAVENGNATWTRDILQCPDDEICTKGEFTPLKATKRNIEVEVLLDKGA 360
Qy 361 KVSADVKKGDPTLHAIIRGRSRKLAELLRNPKDGLYRPNKAGCTPNIDCSHQKSL 420
Db 361 KVSADVKKGDPTLHAIIRGRSRKLAELLRNPKDGLYRPNKAGCTPNIDCSHQKSL 420
Qy 421 TQIFGARHLSPETEDDMLGYDLYSSALADILSEPTMOPPICVGIYAQWGSKSFLLKTL 480
Db 421 TQIFGARHLSPETEDDMLGYDLYSSALADILSEPTMOPPICVGIYAQWGSKSFLLKTL 480
Qy 481 EDEMKTFFAGOOTLEPLQFESWLVFLLLCGGIGLFAFTVHPNIGIANVLSFLALYTF 540
Db 481 EDEMKTFFAGOOTLEPLQFESWLVFLLLCGGIGLFAFTVHPNIGIANVLSFLALYTF 540
Qy 541 FIYIYFGRREGESWMAWVLSRLARHIGYLELLKLMFVNPPELPEQTTKALPYRFLF 600
Db 541 FIYIYFGRREGESWMAWVLSRLARHIGYLELLKLMFVNPPELPEQTTKALPYRFLF 600
Qy 601 TDYNRLSVYGETSLAEMTATLSDACERFGLATRLFRVETEDTQKKKKKTCCLPS 660
Db 601 TDYNRLSVYGETSLAEMTATLSDACERFGLATRLFRVETEDTQKKKKKTCCLPS 660
Qy 661 FVFLFLLIGIISGITLLAIFRVDPKHLTVNAVLISIAVGLAPLVNCRTWQVYLDL 720
Db 661 FVFLFLLIGIISGITLLAIFRVDPKHLTVNAVLISIAVGLAPLVNCRTWQVYLDL 720
Qy 721 NSQKRLHNAASKLHKLKSEGFVKLCVEELMARAKTIDSEFTQNGOTLAVYIIGLDAC 780
Db 721 NSQKRLHNAASKLHKLKSEGFVKLCVEELMARAKTIDSEFTQNGOTLAVYIIGLDAC 780
Qy 781 EODKVLQMDLYAVLEFSKGPFTAIRASDPHIIKAIQNLNSVLRDSNNGHDYMRNVH 840
Db 781 EODKVLQMDLYAVLEFSKGPFTAIRASDPHIIKAIQNLNSVLRDSNNGHDYMRNVH 840
Qy 841 LPEFLNSRGLSNARKFLVTSATNGDVPSCDTTGIOEDARVYSONSLGEMTKLGSSTALN 900
Db 841 LPEFLNSRGLSNARKFLVTSATNGDVPSCDTTGIOEDARVYSONSLGEMTKLGSSTALN 900
Qy 901 RRDYRROMQRTITRQMSFDLTKLVTEDWESDISPQTMRLNLVSYTGRLLRANQIS 960
Db 901 RRDYRROMQRTITRQMSFDLTKLVTEDWESDISPQTMRLNLVSYTGRLLRANQIS 960
Qy 961 FEMDRLASWINLTEQMPYRTSWLILEETEGIPDOMTKTIYERISKNIPTTKDVEPLL 1020
Db 961 FEMDRLASWINLTEQMPYRTSWLILEETEGIPDOMTKTIYERISKNIPTTKDVEPLL 1020
Qy 1021 EIDGDIRNEFVFLSSTPVLAVADYVFLPCTVNDPKLREITIAVRAAREDISIGGLAY 1080
Db 1021 EIDGDIRNEFVFLSSTPVLAVADYVFLPCTVNDPKLREITIAVRAAREDISIGGLAY 1080
Qy 1081 PPLPLEHGGPRAPSGTSPVSCSTSFNGPFAGVVSPQHSYSSGATGPHFFYNG 1140
Db 1081 PPLPLEHGGPRAPSGTSPVSCSTSFNGPFAGVVSPQHSYSSGATGPHFFYNG 1140
Qy 1141 SCGAPGPVYVLLNSLNDAYCEKIQEGIDQSMLEPOYCTTIKKNINGVLAQCNIDEK 1200
Db 1141 SCGAPGPVYVLLNSLNDAYCEKIQEGIDQSMLEPOYCTTIKKNINGVLAQCNIDEK 1200
Qy 1201 KEMNNMFGMHLFRSTVLEMRNAESHVPEDEPRFSESSSGAPHEPRARRASHNLEPHT 1260
Db 1201 KEMNNMFGMHLFRSTVLEMRNAESHVPEDEPRFSESSSGAPHEPRARRASHNLEPHT 1260
Qy 1261 ELSSQPTPLNLSFEELNLTGLDEGAPRHSNLSWOSQTRRTPSLSSINQDSIETSKLT 1320
Db 1261 ELSSQPTPLNLSFEELNLTGLDEGAPRHSNLSWOSQTRRTPSLSSINQDSIETSKLT 1320
Qy 1321 DKVOAETRAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYWGQSSSGSISNLEQEK 1380

```

```

Db 1321 DKVOAETRAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYWGQSSSGSISNLEQEK 1380
Qy 1381 KDSEKPDGGRKSPFLMKRGDVIDYSSGVSNDASPLDPTREDEKSDQSGSKLLPCKKS 1440
Db 1381 KDSEKPDGGRKSPFLMKRGDVIDYSSGVSNDASPLDPTREDEKSDQSGSKLLPCKKS 1440
Qy 1441 SERSLFQTDKLKSGSLRYOKLPSDEDESGTESDNTPLDKDKRAEKYERVPKSP 1500
Db 1441 SERSLFQTDKLKSGSLRYOKLPSDEDESGTESDNTPLDKDKRAEKYERVPKSP 1500
Qy 1501 EHSAPRTFTKAEYLSDALDKDSDSCVRSSESSPNHSLNEVADDSQLEKANLIE 1560
Db 1501 EHSAPRTFTKAEYLSDALDKDSDSCVRSSESSPNHSLNEVADDSQLEKANLIE 1560
Qy 1561 LEDDSHSGKRGIPHSLSGLDPTIARMSICSEDKSPSEGLASSPEENPACOKAYNL 1620
Db 1561 LEDDSHSGKRGIPHSLSGLDPTIARMSICSEDKSPSEGLASSPEENPACOKAYNL 1620
Qy 1621 NRTPTVTLLNNNSAPANRANONFDEMEGIRTSQVILRPSSSPPTTIONENLKSMTNKR 1680
Db 1621 NRTPTVTLLNNNSAPANRANONFDEMEGIRTSQVILRPSSSPPTTIONENLKSMTNKR 1680
Qy 1681 SORSSYTRLSKDPPELHAAASESTGCEERESIL 1715
Db 1681 SORSSYTRLSKDPPELHAAASESTGCEERESIL 1715

RESULT 4
AA096841
ID AA096841 standard; Protein: 1762 AA.
XX
AC AA096841;
XX
DE 30-JUL-2002 (first entry)
XX
DE Rat kidins220 protein from Genbank AA035185.
XX
KW kidins220; kinase D interacting substrate of 22kDa; cytosolic;
KW neuroprotective; gene therapy; protein kinase D; PKd; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
XX
OS Rattus sp.
XX
PN WO200220786-A2.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2001; 2001WO-GB03977.
XX
PR 06-SEP-2000; 2000US-230449P.
XX
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX
PI Schlavo G, Iglesias T;
XX
DR WPI: 2002-371879/40.
XX
PT Polypeptide kinase D interacting substrate of 220 kDa and
PT polynucleotide, useful for identifying modulators useful in treating
PS cancer and neurodegenerative diseases
XX
PS Disclosure: Fig 18; 228pp; English.
XX

The invention relates to polypeptide comprising a kinase D interacting
substrate of 220 kDa (kidins220) from rat and human or their
fragment, variant or fusion provided that the protein is not the
polypeptide encoded by Genbank Accession No. AB03076. Also included are
a polynucleotide encoding kidins220 provided that polynucleotide is
not any one of the clones corresponding to the 61 Genbank Accession Nos.
given in the specification; an expression/replicable vector comprising
the polynucleotide; a recombinant host cell containing the polynucleotide
or vector; an anti-kidins220 antibody (used in the preparation of
kidins220); an agent (AI) which modulates activity of kidins220,

```





QY 1692 DPELHAASSTGTGCEERESTL 1715  
 1740 DASELH-AASSESTGTGCEERESTL 1762

RESULT 5  
 AA080244  
 ID AA080244 standard; Protein; 1763 AA.  
 AC AA080244;  
 XX  
 XX 30-JUL-2002 (first entry)  
 DE Rat kidneys220 protein.  
 XX  
 XX Kidins220; kinase D interacting substrate of 22kDa; cytostatic;  
 KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;  
 KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.  
 XX  
 OS Rattus sp.  
 XX  
 PN W020020786-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 06-SEP-2001; 2001WO-GB03977.  
 XX  
 PR 06-SEP-2000; 2000US-230449P.  
 XX  
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 PI Schiavo G, Iglesias T;  
 XX  
 DR WPI: 2002-371879/40.  
 DR N-PDB: ABK51214.  
 PT Polypeptide kinase D interacting substrate of 220 kDa and  
 PT polynucleotide, useful for identifying modulators useful in treating  
 PT cancer and neurodegenerative diseases -  
 XX  
 PS Claim 33; Fig 1; 228pp; English.  
 XX  
 XX The invention relates to polypeptide comprising a kinase D interacting  
 CC substrate of 220 kDa (Kidins220) from rat and human or their  
 CC fragment, variant or fusion provided that the protein is not the  
 CC polypeptide encoded by GenBank Accession No. AB033076. Also included are  
 CC a polynucleotide encoding Kidins220 provided that polynucleotide is  
 CC not any one of the clones corresponding to the 61 GenBank Accession Nos.  
 CC given in the specification; an expression/replacable vector comprising  
 CC the polynucleotide; a recombinant host cell containing the polynucleotide  
 CC or vector; an anti-Kidins220 antibody (used in the preparation of  
 CC Kidins220); an agent (AI) which modulates activity of Kidins220;  
 CC protein kinase D (PKD) or interaction between PKD and Kidins220;  
 CC a mutant animal transgenic for Kidins220; the use of an agent capable of  
 CC detecting the expression of Kidins220 gene in manufacture of a diagnostic  
 CC reagent for diagnosing or prognosing cancer or for monitoring the  
 CC progression of cancer in a patient; a compound comprising a moiety that  
 CC selectively binds to Kidins220 or its variant and another moiety; and a  
 CC nucleic acid molecule encoding the compound. The antibody is useful for  
 CC modulating the activity of Kidins220, where the antibody prevents a  
 CC region of Kidins220 interacting from another protein, or reduces the  
 CC ability of Kidins220 to bind to adenosine triphosphate (ATP). The  
 CC Kidins220 polynucleotide is useful for identifying an agent which  
 CC modulates the promoter activity of the polynucleotide, and Kidins220 is  
 CC useful for identifying an agent which modulates activity of Kidins220,  
 CC protein kinase D (PKD) and the interaction between PKD with Kidins220,  
 CC Kidins220 or its polynucleotide is useful in medicine, e.g. using gene  
 CC therapy or for diagnosing or monitoring progression of cancer in a  
 CC patient. The agent is useful in manufacture of medicament, for use in  
 CC treatment of neurodegenerative disease, and the agent which inhibits  
 CC function of Kidins220 gene or its product is useful for treating cancer  
 CC which is glioblastoma multiforme or prostate cancer. The compound is

CC useful for imaging cancer in an individual, and for diagnosing or  
 CC prognosing, and also for treating an individual. The present  
 CC sequence represents Rat Kidins220.

XX SQ Sequence 1763 AA:

Query Match 93.4%; Score 8301; DB 23; Length 1763;  
 Best Local Similarity 90.9%; Pred. No. 0;  
 Matches 1603; Conservative 47; Mismatches 64; Indels 50; Gaps 2;

QY 1 MSVLISQSYINVEEBNIPALKALLECKDKDERNECCQTPMLAEOGNEIYKELIKN 60  
 DB 1 MSVLISQSYINVEEBNIPALKALLECKDKDERNECCQTPMLAEOGNEIYKELIKN 60  
 QY 61 GANCNLEDDNNFTALISAKESGHIVHVELLSGASLEHRDGGWTALMAYCYGRITVY 120  
 DB 61 GANCNLEDDNNFTALISAKESGHIVHVELLSGASLEHRDGGWTALMAYCYGRITVY 120  
 QY 121 ELLSHGANPSVTGLQYSYPIIMAAGRHADIVHLLONGAKVNCSDRYGTPLYMAAR 180  
 DB 121 ELLSHGANPSVTGLQYSYPIIMAAGRHADIVHLLONGAKVNCSDRYGTPLYMAAR 180  
 QY 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGYTOSVKEILKRNPNVLTJDKGNTA 240  
 DB 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGYTOSVKEILKRNPNVLTJDKGNTA 240  
 QY 241 LMIASKEGHEIVQDLDAAGTYVNI PDRSGDVLVLAARGGHEIVRLLQKYADIDIRG 300  
 DB 241 LMIASKEGHEIVQDLDAAGTYVNI PDRSGDVLVLAARGGHEIVRLLQKYADIDIRG 300  
 QY 301 QDNKTALYVAVEKGNATWVRDILQCNPDTEICTKDETPPLKATKRNIEVEILLDKGA 360  
 DB 301 QDNKTALYVAVEKGNATWVRDILQCNPDTEICTKDETPPLKATKRNIEVEILLDKGA 360  
 QY 361 KYSAVDKKCDTPLHAIIRGRSRKLAELLRNPKDGRLLRPKAGETPYNIDCSHQKIL 420  
 DB 361 KYSAVDKKCDTPLHAIIRGRSRKLAELLRNPKDGRLLRPKAGETPYNIDCSHQKIL 420  
 QY 421 TQIFGARHLSPTETGDMGVDLYSSALADILSEPTMOPICVGYLAQWGSKSPFLKKL 480  
 DB 421 TQIFGARHLSPTETGDMGVDLYSSALADILSEPTMOPICVGYLAQWGSKSPFLKKL 480  
 QY 481 EDEMTFAQQLEPFOFSWMLIVFTLLCGGLGLFPAFTVHPNIGIAVSFTALALTYF 540  
 DB 481 EDEMTFAQQLEPFOFSWMLIVFTLLCGGLGLFPAFTVHPNIGIAVSFTALALTYF 540  
 QY 541 FVIYIFGGRBEESNMWAVLSTRLARHIGYELLKLMFVNPPLPQOTTKALPVRFLF 600  
 DB 541 FVIYIFGGRBEESNMWAVLSTRLARHIGYELLKLMFVNPPLPQOTTKALPVRFLF 600  
 QY 601 TTYNRLSSVGGSTLAEMATATSDACERFGLATRLRFRVKTEDTGGKKKKKCCCPSS 660  
 DB 601 TTYNRLSSVGGSTLAEMATATSDACERFGLATRLRFRVKTEDTGGKKKKKCCCPSS 660  
 QY 661 FVYFLITIGCIISITGLLAIFFVDPKHLTVNVALVISIASVVGAFVLCRTMWOYLDL 720  
 DB 661 FVYFLITIGCIISITGLLAIFFVDPKHLTVNVALVISIASVVGAFVLCRTMWOYLDL 720  
 QY 721 NSQRRRLHNAASKLHLKSEGFMYKCEVELAMARMAKTIDSEFTONOTRLVYIIDGLDAC 780  
 DB 721 NSQRRRLHNAASKLHLKSEGFMYKCEVELAMARMAKTIDSEFTONOTRLVYIIDGLDAC 780  
 QY 781 BDDKVLQMDTVRYVFSNGPFAIRASPHITIKAINONLSVLDNSINIGHDWRNIVH 840  
 DB 781 BDDKVLQMDTVRYVFSNGPFAIRASPHITIKAINONLSVLDNSINIGHDWRNIVH 840  
 QY 841 LFPVPLNSRGLSNARFLVYSATNGDVPSCDTTQIOEDADRYSQNSLGEWTKLGSKYALN 900  
 DB 841 LFPVPLNSRGLSNARFLVYSATNGDVPSCDTTQIOEDADRYSQNSLGEWTKLGSKYALN 900  
 QY 901 RRDYTRRROMORTITRQMSFDLTKLLVYEDWFSIDSPOTMRRLNIYVTRGLLRANOIS 960  
 DB 901 RRDYTRRROMORTITRQMSFDLTKLLVYEDWFSIDSPOTMRRLNIYVTRGLLRANOIS 960

```

QY 961 FNMRLASMINITEOMPYRTSNLLIYLETETGIPDQMLTKTYERISKNIPTTKDVEPL 1020
DB 961 FNMRLASMINITEOMPYRTSNLLIYLETETGIPDQMLTKTYERISKNIPTTKDVEPL 1020
QY 1021 EIDGIRNPEVFLSRTPVLAARDVKFLPCTVNDPLRLTIDVRAAROIISGIGLAY 1080
DB 1021 EIDGIRNPEVFLSRTPVLAARDVKFLPCTVNDPLRLTIDVRAAROIISGIGLAY 1080
QY 1081 PPLPLHEGPPRAPSGYSGPPSYCSSTSFNGFPAGVSPQPHSSYSGMTGPCHPFYNR- 1139
DB 1081 PPLPLHEGPPRAPSGYSGPPSYCSSTSFNGFPAGVSPQPHSSYSGMTGPCHPFYNR- 1139
QY 1140 -----GSPAPGPVVL 1151
DB 1141 FFAPLYLPYRYPGGSOHLISRSYKASLPRDQNNGLPCDGSFNGKROGSRPCHROLVT 1200
QY 1152 NSLNDVAVCEKLOJEGDQSMLPQYCTTIKKANINGRYLAQCNDIDELKEMNNFGDMH 1211
DB 1201 EPNDCGCRCEKROJEGDQSMMPQYCTTIKKANINGRYLAQCNDIDELKEMNNFGDMH 1260
QY 1212 LFRSTVLEMRNAHVPEDEPRLISESSSGAPHPGEPARRASHNELPTELSSQTPYTLN 1271
DB 1261 LFRSTVLEMRNAHVPEDEPRLISESSSGAPHPGEPARRASHNELPTELSSQTPYTLN 1320
QY 1272 FSEELNLTGIDEGAPRRHSNLSMOSQTRTPSLSLNSQDSIEISKITDKVOAFYRDAY 1331
DB 1321 FSEELNLTGIDEGAPRRHSNLSMOSQTRTPSLSLNSQDSIEISKITDKVOAFYRDAY 1380
QY 1332 REYIAQMSQLEGGPSTTISGRSPHSTYYMGQSSGSIHNSLEQEKDSEKPDGGR 1391
DB 1381 REYIAQMSQLEGGPSTTISGRSPHSTYYMGQSSGSIHNSLEQEKDSEKPDGGR 1440
QY 1392 KSFLLKRDVDYDISSSYSTNDASPLDITEDEKSDQSGSKLLPGKSSERSLFPQDL 1451
DB 1441 KSFLLKRDVDYDISSSYSTNDASPLDITEDEKSDQSGSKLLPGKSSERSLFPQDL 1500
QY 1452 KFKGGLRQKLPSPDESGTESDNTPLKDDKRAKGAVERYPKSPHSAEPIRTFI 1511
DB 1501 KFKGGLRQKLPSPDESGTESDNTPLKDDKRAKGAVERYPKSPHSAEPIRTFI 1560
QY 1512 KAKEYLDALDKDSDSGVRSSESSPNHSLANEVADDSOLEKANLIELEDDSHSGRG 1571
DB 1561 KAKEYLDALDKDSDSGVRSSESSPNHSLANEVADDSOLEKANLIELEDDSHSGRG 1620
QY 1572 IPHSLSGIOPPIAAMSTICSEDKKSPSECSLASSPEBSWPCQAKAYMLNRPSTVTLN 1631
DB 1621 IPHSLSGIOPPIAAMSTICSEDKKSPSECSLASSPEBSWPCQAKAYMLNRPSTVTLN 1680
QY 1632 NSAPANRANONDEMEGIRETSOVLIRPSSSPNPPTTIONENLKSMTHRSSOYSTRLSK 1691
DB 1681 NSAPANRANONDEMEGIRETSOVLIRPSSSPNPPTTIONENLKSMTHRSSOYSTRLSK 1740
QY 1692 DPPELHAASSESTGFGERESIL 1715
DB 1741 DASELH-AAASSESTGFGERESIL 1763

```

```

RESULT 6
ID AAG75604 standard; Protein; 705 AA.
AC AAG75604;
XX
XX 03-SEP-2001 (first entry)
XX
XX Human colon cancer antigen protein SEQ ID NO:6368.
XX
XX Human colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.
XX Homo sapiens.
XX

```

```

PN W020122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000MO-US26524.
XX
XX 29-SEP-1999; 99US-0157337.
XX 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI, 2001-235357/24.
XX N-PSDB; AAH35009.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX Claim 11; Page 7830-7834; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patient's own production of P.
XX Additionally, N may be used to produce the colon cancer-associated Ps,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAB77789 represent sequences used in the exemplification of the
XX present invention.
XX
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 705 AA:
XX
XX Query Match 34.3%; Score 3050.5; DB 22; Length 705;
XX Best Local Similarity 98.8%; Pred. No. 1e-234; 4; Indels 1; Gaps 1;
XX Matches 587; Conservative 2; Mismatches 4;
XX
XX 62 ANCNLELDLNMWTLISASKEGHVHVEELKCGVNLHRDMGWTALMACYKGRDVE 121
XX :
XX 28 SNCNLELDLNMWTLISASKEGHVHVEELKCGVNLHRDMGWTALMACYKGRDVE 87
XX
XX 122 LLLSHGNPSPVTLQSVYPIIWAAGRGADIVHLLONGAKVNCSDKGTPIVAAAK 181
XX
XX 88 LLLSHGNPSPVTLQSVYPIIWAAGRGADIVHLLONGAKVNCSDKGTPIVAAAK 146
XX
XX 182 GHLECYKHLAMGADYDQEGANSMTALIYAVKGYTOSYKEILKRNPNVNLTDKDGNTAL 241
XX
XX 147 GHLECYKHLAMGADYDQEGANSMTALIYAVKGYTOSYKEILKRNPNVNLTDKDGNTAL 206
XX
XX 242 MIASKEGHEIYODLLDAGTYVNIIPRSGDTVLIGAVRGHVEIVALLQKADIDIRQ 301
XX
XX 207 MIASKEGHEIYODLLDAGTYVNIIPRSGDTVLIGAVRGHVEIVALLQKADIDIRQ 266
XX
XX 302 DNKFTALYMAEKNAMVNRDILQCNPDTECTFQDGEPLIKATKMNIEVVELLDKGAK 361
XX
XX 267 DNKFTALYMAEKNAMVNRDILQCNPDTECTFQDGEPLIKATKMNIEVVELLDKGAK 326
XX
XX 362 VSAVDDKGDPTLAIATRGSRKLAELLRNPKDGRLLYPPNKAGETPYNIDCSHOKSLIT 421
XX
XX 327 VSAVDDKGDPTLAIATRGSRKLAELLRNPKDGRLLYPPNKAGETPYNIDCSHOKSLIT 386
XX
XX 422 QIFGARHLSPTEEDGMDLGYDIYSSALADILSEPTMOPPICVGLAQMWSGKSFLLKLE 481
XX

```



DB 420 HNNLGLAVSLFLALITYITVITGGRNREGSMWAWLSTRLARHIGYELLLKLMFV 479  
 QY 582 NPPELPEQTTKALPVRFLETDYNNRLSVGGESTLAEMTATLSDACEREGFLATRLFRVF 641  
 DB 480 NPPELPEQTTKALPVRFLETDYNNRLSVGGESTLAEMTATLSDACEREGFLATRLFRVF 539  
 QY 642 KTEEDTQGGKKKK 653  
 DB 540 KTEEDTQGGKKKK 551

RESULT 8  
 AAB95191  
 ID AAB95191 standard; Protein; 543 AA.  
 AC AAB95191;  
 XX 26-JUN-2001 (first entry)  
 DT  
 XX Human protein sequence SEQ ID NO:17273.  
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 OS Homo sapiens.  
 XX  
 XX EPI074617-A2.  
 PN 07-FEB-2001.  
 PD  
 XX 28-JUL-2000; 2000EP-0116126.  
 PF  
 XX 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 PT  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 XX Claim 8; SEQ ID 17273; 2537pp + CD ROM, English.

CC of the present invention.  
 XX  
 SQ Sequence 543 AA;  
 Query Match 31.4%; Score 2791; DB 22; Length 543;  
 Best Local Similarity 99.6%; Pred. No. 3.9e-214;  
 Matches 541; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1173 MLPQYCTTKKANINGVLAQCNIDELKRMNNFGDMHLEFRSTVEMRNASHVVEP 1232  
 DB 1 MLPQYCTTKKANINGVLAQCNIDELKRMNNFGDMHLEFRSTVEMRNASHVVEP 60  
 QY 1233 RFLSESSSGPAPGEPARRASHNELPHTLELSQPTPLNFSFEELNTLGLDEGAPRHSNL 1292  
 DB 61 RFLSESSSGPAPGEPARRASHNELPHTLELSQPTPLNFSFEELNTLGLDEGAPRHSNL 120  
 QY 1293 SWSQTRRTPSLSLNSQDSSTISKTQVQAEYADAYREYIAQMSOLEGGPGSTTISG 1352  
 DB 121 SWSQTRRTPSLSLNSQDSSTISKTQVQAEYADAYREYIAQMSOLEGGPGSTTISG 180  
 QY 1353 RSSPHSTYYMGQSSGSIHSNLEOKRGKDEPKPDGKRKSPFLMRGVDYIDSSGCVSTN 1412  
 DB 181 RSSPHSTYYMGQSSGSIHSNLEOKRGKDEPKPDGKRKSPFLMRGVDYIDSSGCVSTN 240  
 QY 1413 DASPLDPTFEDEKSDQSGSKLLPGKSSERSSLFQTDILKSGSLRYQKLPDSDESGT 1472  
 DB 241 DASPLDPTFEDEKSDQSGSKLLPGKSSERSSLFQTDILKSGSLRYQKLPDSDESGT 300  
 QY 1473 EESDNTPLKDDDKRAEKGVERVPRKSPESAEPITFTIKAEYISDALDKKSDSGV 1532  
 DB 301 EESDNTPLKDDDKRAEKGVERVPRKSPESAEPITFTIKAEYISDALDKKSDSGV 360  
 QY 1533 RSESSSPNHSIHNEVADDSOLEKANLIELEDSSHSGRGIPLSLGLDPIIARMSICSE 1592  
 DB 361 RSESSSPNHSIHNEVADDSOLEKANLIELEDSSHSGRGIPLSLGLDPIIARMSICSE 420  
 QY 1593 DKSPSECSLIASSPENNPACOKAVNLRFTSTVTLNNSAPANRANONFDEMGIRET 1652  
 DB 421 DKSPSECSLIASSPENNPACOKAVNLRFTSTVTLNNSAPANRANONFDEMGIRET 480  
 QY 1653 SOYLTPSSPNPTTONTNLSKMTKRSORSSTYPLSDPELHAASSSESTGGERE 1712  
 DB 481 SOYLTPSSPNPTTONTNLSKMTKRSORSSTYPLSDPELHAASSSESTGGERE 540  
 QY 1713 SIL 1715  
 DB 541 SIL 543

RESULT 9  
 AAE03645  
 ID AAE03645 standard; Protein; 513 AA.  
 XX  
 AC AAE03645;  
 DT  
 XX 06-AUG-2001 (first entry)  
 DE  
 XX Human extracellular matrix and cell adhesion molecule-9 (XMA-9).

Human; extracellular matrix and cell adhesion molecule: XMA-9;  
 KW gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;  
 KW Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;  
 KW sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;  
 KW inflammatory disorder; acquired immune deficiency syndrome; AIDS;  
 KW Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;  
 KW Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;  
 KW glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;  
 KW osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;  
 KW infection; cell proliferative disorder; actinic keratosis; myeloma;  
 KW arteriosclerosis; neotrophic; anticonvulsant; antithyroid; nephrotropic;  
 KW neuroprotective; dermatological.  
 XX  
 OS Homo sapiens.

```

XX Key Location/Qualifiers
FH Region 94..108
FT Domain /note="Adrenocorticotrophin receptor"
FT Domain 196..198
FT Region /note="Rgd cell interaction motif"
FT Region 328..335
FT /note="Phage integrase family"
XX
XX W0200142285-A2.
XX
XX 14-JUN-2001.
XX
XX 05-DEC-2000; 2000MO-US32990.
XX
XX 10-DEC-1999; 9905-0172852.
XX 16-DEC-1999; 9905-0172354.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H. Tang YT, Lal P, Burford N, Azimzai Y, Patterson C,
XX Baughn MR, Lu DM, Shah P, Au-Young J;
XX
XX WPI; 2001-381632/40.
XX
XX N-PSDB; AAD08053.
XX
XX New human extracellular matrix and cell adhesion molecules and
XX polynucleotide sequences encoding them, useful for diagnosis,
XX prevention, treatment of genetic, autoimmune and cell proliferative
XX disorders
XX
XX Claim 1; Page 101-102; 135pp; English.
XX
XX The present sequence is a human extracellular matrix and cell
XX adhesion molecule (XMAP). The XMAP is used for screening a compound for
XX effectiveness as an agonist or antagonist of XMAP. The identified agonist
XX or antagonist are used for treating a disease or condition associated
XX with decreased or increased expression of functional XMAP. The
XX polynucleotides encoding XMAP are useful in somatic or germline gene
XX therapy to correct a genetic deficiency, to express a conditionally
XX lethal gene product and to express a protein which affords protection
XX against intracellular parasites and also for diagnosis of disorders
XX associated with expression of XMAP. They are also used for generating
XX hybridisation probes useful in mapping the naturally occurring genomic
XX sequences and to create knock in humanised animals (pigs) or transgenic
XX animals (mice or rats) to model human diseases. Oligonucleotide or longer
XX fragments derived from the polynucleotide sequences may be used as
XX elements on a microarray. Antibodies which specifically bind XMAP may be
XX used for the diagnosis of disorders associated with the expression of
XX XMAP, or in assays to monitor patients being treated with XMAP. Diseases
XX diagnosed, prevented or treated include genetic disorders such as
XX adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
XX disease, myotonic dystrophy, sickle cell anaemia, thalassemia,
XX autoimmune/inflammatory disorders such as acquired immune deficiency
XX syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
XX atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
XX glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
XX osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
XX bacterial, fungal, parasitic, protozoal and helminthic infections and
XX cell proliferative disorders such as actinic keratosis, arteriosclerosis
XX and cancer including breast, bladder, bone marrow, brain and uterus
XX cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
XX
XX Sequence 513 AA;
XX
XX Query Match 29.7%; Score 2638; DB 22; Length 513;
XX Best Local Similarity 99.8%; Pred. No. 6.3e-202;
XX Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1203 MNMFGDWHLFRSTVYLEMRNESHVPEDEPFLSESSSGPAPHPGEPARRASHNELPTEL 1262
XX 1 MNMFGDWHLFRSTVYLEMRNESHVPEDEPFLSESSSGPAPHPGEPARRASHNELPTEL 60

```

```

QY 1263 SSQTPYTLNFSPEELNTGLDGCAPRHSNLSWQSOTRRTPSLSLNSDSSIEISKLTDK 1322
DB 61 SSQTPYTLNFSPEELNTGLDGCAPRHSNLSWQSOTRRTPSLSLNSDSSIEISKLTDK 120
QY 1323 VOAEYRDAYREYIAOMSLEGPGSTTISGRSSPHSTYTYMGQSSGSGSIHSLNDEKGD 1362
DB 121 VOAEYRDAYREYIAOMSLEGPGSTTISGRSSPHSTYTYMGQSSGSGSIHSLNDEKGD 180
QY 1383 SEPKPDGKRSFLMRGVDIDYSSGVSFNDAISPDPITPEDEKSDQSKLPEKKSSE 1442
DB 181 SEPKPDGKRSFLMRGVDIDYSSGVSFNDAISPDPITPEDEKSDQSKLPEKKSSE 240
QY 1443 RSLFOTDLKLGSGLRQKLPSPDESEGTSESDNTPLKDKDKKAKGKYERPKSPDH 1502
DB 241 RSLFOTDLKLGSGLRQKLPSPDESEGTSESDNTPLKDKDKKAKGKYERPKSPDH 300
QY 1503 SAEPITFTKAEYISDALIDKSDSDGVSRSSESPNHSILNEYVADSOLEKANLILE 1562
DB 301 SAEPITFTKAEYISDALIDKSDSDGVSRSSESPNHSILNEYVADSOLEKANLILE 360
QY 1563 DSHSGKRGIPHSLSGLDPIIARMSICSEDKKSPSECSLIASPEENWPACOKAYNLNR 1622
DB 361 DSHSGKRGIPHSLSGLDPIIARMSICSEDKKSPSECSLIASPEENWPACOKAYNLNR 420
QY 1623 TPSTVTLNNSAPANRANONFDEMGIRETSOYLIRPSSSPPTTIONENLKSMTHKRSQ 1682
DB 421 TPSTVTLNNSAPANRANONFDEMGIRETSOYLIRPSSSPPTTIONENLKSMTHKRSQ 480
QY 1683 RSTYTLSDPPELHAAASSESTGGEERESTL 1715
DB 481 RSTYTLSDPPELHAAASSESTGGEERESTL 513

RESULT 10
ABG08697
ID ABG08697 standard; Protein; 624 AA.
XX
XX ABG08697;
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #8688.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX W0200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001MO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Dmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS72884.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX Claim 20; SEQ ID NO 39056; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

```







```

QY 420 LTOIGARHLSPETDGMGLDYLYSSALADILSEPTMOPICGVGYAOMSGSKSFLKK 479
DB 442 LGQVFGARLINTNDESGMLYELYSALADVLSEPTLTPTIVGLYAKMGSKSFLKK 501
QY 480 LEDEMKTFAGQOIEPLFQFSMLVFLTL---LLCG---GGLLFAFTVHNLGIANVLSF 533
DB 502 LREMNNAFARQMAEPPIRTSGLLFVCLHVALLIGIVGISTWSAV---VGVSAVGF 556
QY 534 LALLYFIVYFGREGESGMMWAWLSTRARHIGYELLKLMFVNPELPEQOTTKA 593
DB 557 LLLAYLLAARVCNVO---MDQMAYSVOGLEKRMRLMLLIQVAFCHPPG---PQSDQA 613
QY 594 LPVAFLEFDYNRSLSSVGGETSLAEMITLSDACEREFGLATRLFRVFKTE---DTQKKR 651
DB 614 KPVAFHFAEANSASPT---GDGAVAHMLAALLDAIESHGMATRLRYAFRCKLKVGYMR 672
QY 652 MKKTCCLPSFVIFLITIGIISGTTLAIF---RDPKHLVYNAVLIASVGLAFV 706
DB 673 WRRKCCIPYLIFELALVTVVGTISLVAFTFADEKEKHILV---ALVYIAVMGTLIC 730
QY 707 LNCRTMVOVLDLNLISQKRRLHNAASKRLKSEGFVKLCVELMARAKTIDSPTON 766
DB 731 THLHVLAKVVSLETSIRVLKRAV---RSSASAPLTMGAIVAVMTMVKCLDAFTNQ 786
QY 767 QTRLVLIIDGLDACEODKVLQMLDYRVAFS---KGPIALFASDPHIIIRAINQNSVL 824
DB 787 QSRIVGYIDALDSCDERITLLTNAVQTLSSPNRPVLLISDPHVIKAAABANSRRLF 846
QY 825 RDSININGDMRNIVHLPVFLNRSGLSNARKFLVTS---AINGVPCSD---TTGQED 877
DB 847 TEGIGGHDFLRNLVHLPVYLQNSGLKRVQRAQMTALLFRSGGDDQTDGPTLIGSVS 906
QY 878 ADDRVSONS---LGEMLTL---GSKTALN---RDTYRRMQRTITROMSF 920
DB 907 A-RRLSNASELISQEKLRGPARGGGKRLISESVASGSLNHLRGMPQVY---L 960
QY 921 DLKRLVTEWFSDDISQYMRRLINIVSYGRLIRANQISFNDRLASWINTLEQWYRT 960
DB 961 DLSRIVLTIDYFSDVNRSMRRLMNVIVYIRVRLKAFQIEFSWYRLSSWINTLEQWPLRA 1020
QY 981 SWILYILEE---TEGIPOMTLKTYERISKINIPITTKOVBELLEDGDIRFEVFLSRT 1038
DB 1021 SMYLNHDQWMSNADSVSLQSYTEKLRKLYARLREAPLLEDRERKLDLAFLOHKS 1080
QY 1039 VLVAADKVELPCTVNDLPKRLRETIAD---VRAREQISIGGLAAPPPLRHE 1087
DB 1081 DLVAVDLRIFLPFTINLDPLRYKLAKEDQCTIEDEGLVIAQARPSVNTWRQRP----- 1134
QY 1088 GPPRABSGYQSPSVCSSTFNGPFAGVVSPQPHSSY-----YSGMTG 1131
DB 1135 ---APPTY---VSPQAYPYQMFONEYPANELRSNLSSTSTE 1171
QY 1132 POHPFYRSGPAPGAPVY---LNSLNDVAVCEKIKOJEGDQSLPQCTTIKANINGR 1189
DB 1172 PVPPLINSPDSREDLQTLKLDLYEGVYISLDRJEDM-KALPKLAVLYLENAINGR 1230
QY 1190 VLAQCNIDELKKEKMNNGDMHLEFRSTVLEMRNAESHVVP-----EDPRFLESSESGP 1242
DB 1231 VLKHCMDPLKSVLGSFGHWELEFRLLITTLREOER-LPRKQOQQQCGALAEAPSNVP 1288
QY 1243 APHG-----EPARR---ASHNELPHE---LSSQ---TPYLLNLSFEELNLTGLDEGA 1286
DB 1289 MIKDVTDALMQPPRESLSRKNNSVSHMEKQVTEOMICGTQLINEAIVE-DVAASSERPS 1347
QY 1287 PRHNLNLSMOSOTRPTPLSL-----NSODSIEISKILTDVQ---AAY-EDAYREYI 1335
DB 1348 PTGMLAIVADOLAPLRESSESGSPDDQKQYGVKLSNNNNNOYTHAETNNSVSSHSL 1407
QY 1336 AOMSOLEGPGSTTISGRSSPHSTTYMGQSSSGSISHLNLEOEKGD 1382
DB 1408 QSLSTLVGAP-----VGHGSGSG---SHHLTGND 1435

```

```

RESULT 12
AAB94785
ID AAB94785 standard; Protein; 129 AA.
AC AAB94785;
DT 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:15890.
XX
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000BP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELT-) HELIX RES INST.
PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 15890; 2537bp + CD ROW; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 129 AA;

```

Query Match 7.5%; Score 666; DB 22; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-45;  
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1587 MSICSEDKKSPSCSLIASSPENNPPACQKAYNLNRPTSTVTLNNSAPANRANONPEM 1646
DB 1 MSICSEDKKSPSCSLIASSPENNPPACQKAYNLNRPTSTVTLNNSAPANRANONPEM 60

```

QY 1647 EGIRETSQVILRPSSSPNPTTTONENLKMTHKRSORSSYTRLSKDPPELHAAASSESTG 1706  
 |||||||  
 Db 61 EGIRETSQVILRPSSSPNPTTTONENLKMTHKRSORSSYTRLSKDPPELHAAASSESTG 120  
 QY 1707 FGEERESIL 1715  
 |||||||  
 Db 121 FGEERESIL 129

RESULT 13  
 AAM40811  
 ID AAM40811 standard; Protein: 120 AA.  
 AC AAM40811;  
 XX 22-OCT-2001 (first entry)  
 DT Human polypeptide SEQ ID NO 5742.  
 DE Human; nocrotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia.  
 XX Homo sapiens.  
 OS WO200153312-A1.  
 XX 26-JUL-2001.  
 PD 26-DEC-2000; 2000WO-US34263.  
 PF 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX (HYSE-) HYSEQ INC.  
 PA Tang YR, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPI: 2001-44253/47.  
 XX N-PSDB: AAI59967.  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 XX such as central nervous system injuries -  
 PS Example 2; SEQ ID NO 5742; 10078pp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nocrotropic.  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localized neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX SO Sequence 120 AA;  
 Query Match 7.0%; Score 620; DB 22; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-41;  
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 TOSVKEITLKRPNVNLTKDKNFALMIASKEGTEIVODLDAAGTVNIPDSGGPVLLG 276  
 |||||||  
 Db 1 TOSVKEITLKRPNVNLTKDKNFALMIASKEGTEIVODLDAAGTVNIPDSGGPVLLG 60

QY 277 AVRGHVEIYRALLQKADIDIRGODNKATLYAVKGNATVYRDILQCNPTETCTKDG 336  
 |||||||  
 Db 61 AVRGHVEIYRALLQKADIDIRGODNKATLYAVKGNATVYRDILQCNPTETCTKDG 120

RESULT 14  
 AAB56277  
 ID AAB56277 standard; Protein: 109 AA.  
 AC AAB56277;  
 XX 13-MAR-2001 (first entry)  
 DT Human secreted protein sequence encoded by gene 82 SEQ ID NO:371.  
 DE Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 KM antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 KM cerebroprotective; nocrotropic; neuroprotective; antibacterial; virucide;  
 KM fungicide; ophthalmological; gene therapy; pathological condition;  
 KM autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;  
 KM neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischemia;  
 KM cerebrovascular disorder; angiogenesis; nervous system disorder;  
 KM Alzheimer's disease; infection; ocular disorder; corneal infection;  
 KM wound healing; skin aging; food additive; preservative.  
 XX Homo sapiens.  
 OS WO200070042-A1.  
 XX 23-NOV-2000.  
 PD 11-MAY-2000; 2000WO-US12788.  
 PF 13-MAY-1999; 99US-0134068.  
 PR (HUMAN-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;  
 PI Duan RD, Florence KA, Soppet DR;  
 XX WPI: 2000-679828/66.  
 PT Isolated nucleic acid molecule encoding a human secreted protein is  
 XX used in preventing, treating or ameliorating a medical condition -  
 PS Disclosure; Page 1025-1026; 1065pp; English.

The polynucleotide sequences given in AAC999818 to AAC99977 encode the  
 CC human secreted proteins given in AAB56077 to AAB56362. Human secreted  
 CC proteins have activities based on the tissues and cells the genes are  
 CC expressed in. Examples of activities include: immunosuppressive;  
 CC antirheumatic; antiproliferative; cytostatic; cardiant;  
 CC vasotropic; cerebroprotective; nocrotropic; neuroprotective; antibacterial;  
 CC virucide; fungicide; and ophthalmological. The human secreted  
 CC polynucleotides and proteins can be used to prevent, treat or ameliorate  
 CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,  
 CC dogs, chickens or sheep. They are also used in diagnosing a pathological  
 CC condition or susceptibility to a pathological condition. Disorders which  
 CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid  
 CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or  
 CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular  
 CC disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders

CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
 CC fungi and ocular disorders e.g. corneal infection. The proteins can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a  
 CC food additive or preservative to increase or decrease storage  
 CC capabilities. AAC99809 to AAC99817 and AAB56076 represent sequences used  
 CC in the exemplification of the present invention.

XX Sequence 109 AA;

Query Match 6.28; Score 553; DB 21; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-36;

Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 DNKTALYWAWEKGNATVWRDILQCNPTETCTKDGEPPLKATKMRNIEVELLDKGA 361

DB 1 DNKTALYWAWEKGNATVWRDILQCNPTETCTKDGEPPLKATKMRNIEVELLDKGA 60

QY 362 VSAVDKKGDTPLHAIIRGSRKLAELLRNPKDGRLLRPNKAGETP 408

DB 61 VSAVDKKGDTPLHAIIRGSRKLAELLRNPKDGRLLRPNKAGETP 107

RESULT 15  
 AAM79160  
 ID AAM79160 standard; Protein; 1872 AA.

AC AAM79160;

DT 06-NOV-2001 (first entry)

DE Human protein SPQ ID NO 1822.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

OS nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

PN W0200157190-A2.

PD 09-AUG-2001.

XX 05-FEB-2001; 2001MO-US04098.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0653325.

PR 30-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX WPI; 2001-476283/51.  
 DR N-PSDB; AAK52293.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20; Page 4199-4202; 6221pp; English.  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK60302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC actin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM6020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX Sequence 1872 AA;

Query Match 5.98; Score 526; DB 22; Length 1872;  
 Best Local Similarity 20.6%; Pred. No. 8.3e-32;

Matches 423; Conservative 271; Mismatches 738; Indels 618; Gaps 80;

QY 14 EENIPALKALEKCKDVERNECOTPLMAEONGLEIYELIKNGANCLIEDJMT 73

DB 73 KEGHGLVDELGRSSVDSATKGNALHIASLAGAEVAVLVKGANINIAOSONGFT 132

QY 74 ALISAKKEGHVHIVEELKGVN-----LEHRDMG--- 103

DB 133 PLYMAAQNENHIDVYVYLLENGANOSTATEDGFTPLAVALQGHNAVALLENDTKGKVR 192

QY 104 -----GWTALMAACYGRDVELL 123

DB 193 LPALHIAARKDDKTSALLLQNDHNADYOSKMMVNRTESEFTPLHIAHGNVAVDIL 252

QY 124 LSHGANPSVTGLQYSVYPIIWAAGRGHADIYHLLQNGAKVCSQYCTTPLVWAARG- 182

DB 253 LNRGAADVETA-RNGITPLHVASKRGNTMVALLDREGQIDAKTRDGLTPHCAARGH 311

QY 183 -----HLECVHLLAMGADVDOEGANSWTALIV 210

DB 312 DQVELLEERGAFLARTKNGLSPLHMAAQGBHVCVNHQHKARPVDVLDLYTALHY 371

QY 211 AVKGYTQSVKEILKRNPNVNLTDQNTALMASKE----- 247

DB 372 AACHGHRYVTKLLDCKRANPNARALNGFTPLHIAKCKNRKIKMELLYKASIQATRESG 431

QY 248 -----GTEIYODLLDAGTYVNIIPDRSGDVLFGAVRGHVEIVRALCKYADID 297

DB 432 LTPHVAAPMGLNIVLLILQNGASPDVTNIRGETALHMAARAGQVEYVCKLRGALVD 491

QY 298 IRGDNKTALYWAWEKGNATVWRDILQCNPTETCTKDGEPPLKATKMRNIEVELLD 357

DB 492 AARAEQPLHIASRLGKTEIVQILLQMAHBDATNGYTPHISAREGQDVASVLE 551

QY 356 KKAQVSAVDKDDTPLHAIIRGSRKLAELLR-----NPKD 394

DB 552 AGAASHLATKGFPLHVAAKYGLDVAKLLQRRRAADSAGKNGPLHVAAYDOKV 611

QY 395 GRLL-----YRPKAGETPYNIDC-SHOKSILROI--FGAR-----HLSP 432

DB 612 ALLILEKASPHATKNGYTPPLHIAAKKNQOIASTLNTYGAENITYTKQGYTPHIASQ 671

QY 433 ETDGDM-----GYDLYSSALADILSEPTMQPPICVGLYAWQWSGSKFLKLEDEM-- 484

DB 672 BGHTDMVTLDDKGANIHMSKSGITS-----LHLAQ-----EDKYNV 710

QY 485 -----KTRAGQOIEPLFQSMILYVLTLLC--GGGLL--FAFYVHPMLGIAVISLAL 536

DB 711 ADILTKHGADDOAHHTKIGYTPPLV-----ACHYGVKVMNPLIKOGANVNAKTKNGYTP 765

QY 537 -----LYFEIYIFCGRREGESNMAMWLSTRLAHNGYLELLKLMFVNP-- 584

DB 766 HQAAQGGTHITNLVLOHGAKPNATYANGNTALA--IAKRGYISYVDTLKYVTEEVYTT 823

QY 585 -----ELPEQTTALPVRLEFTDYNRLSSVSGETSIAEMIATLSDACEREFGLA 634

Db 824 TTTTTEKHLNVPETMEFVLDVSDDEGD-DTMTGDGGEYLPEDLKELDSDSSQFLD 882  
 QY 635 TRLRVEFTEDTQGGKKKKKCTCCPBYVIFLFIIGCIIISGTTLLAIRVDPKHLTVAVL 694  
 Db 883 GMMNLRYSLEGGSRDS-----LRSSSDSHLSHASY 915  
 QY 695 ISIASVGLAVVLCRTWMOY---LDSLNSOR---KRLHAA---SKLHK----- 736  
 Db 916 LRDSVMDSVYIYSH---QYSTLAKERNSYLSMGTENDVALSSSPHSGFLVSP 972  
 QY 737 -----LKSEGFVKVL--KC--EVELMARMAKTIDSTONOTRLVY---IIDGL 777  
 Db 973 MYDARGAMRGCRHNGHIIIPPRKCTAPTATVTCRLVK-----RHRLATMPMEGE 1024  
 QY 778 DACEODKVLQMLDYRVLFSGPFIATFASDPH--IITAINQNLNSYLRSNIN--GHD 833  
 Db 1025 GLASR-----LIEVPGSAQOFLGKHLPTAPPLNEGESVLSRLIQGPCTK 1072  
 QY 834 YMRN-IYHLPVFLNSRGLSNARKFLVTSATNGDVPCSDTGTQEDADRVVSQNSLGEMTK 892  
 Db 1073 FLGPVIVFIFHPALRG--KERELVYLSENGDSMKHEFCDTED-----ELNELIN 1122  
 QY 893 LGSKTALNRDPTRRRQMOFTITROMSFDTLKLVTEDWFSDISPQIMRLNITVSVGR 952  
 Db 1123 -GMDEVLDSPEDLEKRICRITITRDE-----POYFAVVS--RIKODSNLIGPEG 1169  
 QY 953 LLRAN-----QISFWMDLASMINTEQWPRYRTSMILLYLETEGIPQOMTKTIVRIS 1007  
 Db 1170 VLSSTVVPQOVAVPEGALTKRIRY-----GLQAQPMHSELVKIL 1210  
 QY 1008 KNIPPTKDEPLELEIDGIRNFEVFLSSRTPLVAVADVKVLPCTVNLDPKIREIADVR 1067  
 Db 1211 GNKAT---FSPIVLTLEPRRRKFKPITWTIPV-----PK----- 1241  
 QY 1068 AAREGISIGCLA--YPLPLHGGPRAPSGYSQPPVCSSTSENGPRAGVVSQPPHSSY 1125  
 Db 1242 -ASSDVMLNGFGGDAFPLRL-----LCSIT-----GGTTPAQ----- 1272  
 QY 1126 YSGMTGPQHPFYNRGSPAPGPVYLLN-----SLNVDA-----VCEKLOI----- 1166  
 Db 1273 WEDITGT-----PLTFVNECVSTTNVSARFMLIDCRQIQESVTFASQVYR 1319  
 QY 1167 BGLDQMLPQYCTTITKAN-INGRYLAQCNIID--ELKEMNMNGDWHLEFRSTVLEMRN 1222  
 Db 1320 EITCVYAKKFVFAKSHDPIEARLCFCMTDDKVDKTLBOQENPAE--VARSRDVEYLE 1377  
 QY 1223 AESHYVPEDPRFLSSSSGPARPGEPARASHNELP---HTEISSQTPY-TLNSFPEELN 1278  
 Db 1378 KRPVYVDFGMLVPLTKSG-OHHLFSFPAFKENRLPLFKVVRDVTQEPGRLSPMKPEKS 1436  
 QY 1279 TGLDEGAPRHSNLSWQSOTRTPSLNSLNSODSSIETSKLDKVOAERYDAYREYIAQM 1338  
 Db 1437 TRGLVHQAICNINILPTITKES-----DOBOEEIDMTSEKNPDQERIEERLAYI 1491  
 QY 1339 SQLEGGPSTTISGR-----SSPHSTYYMGQSSSGSIHNSLDEQKGDSEPPKDDGRK 1392  
 Db 1492 ADHLG--FSWTELARELDFTEBOIHQRIENPNSLODQSHALKYLERDGHATDTVLV 1549  
 QY 1393 SFIMK--RGDVIDYSSSVTNDASPLDPTIEDEKSDQSSKLLPGKSSERSSLFOTD 1450  
 Db 1550 ECLTQIRHMDIVHL---METN---TEPIQERISHSYAEIEQIT--LDHSEGFSLQEE 1600  
 QY 1451 LKLKSGGLRYOKLPDEDESGTEESDNTPLKD-----DKDKAKGKVERVPKS 1499  
 Db 1601 LCTAOKROKEROAVSKESET-----CDHPPIVSEEDISVGYSTFODGVKXTGDDSSSTALF 1656  
 QY 1500 PEHSAEPIRTFIKAK-----EYLSDALLDKDDSDSGVRSSE-----SSPNH 1541  
 Db 1657 PQTKEQYVOQDFSGMQDLPRESSLEYQOEYFV-----TTPGTETSETQKAMIVSSPSK 1711  
 QY 1542 SLHNRYADDSQLEKANLIELEDDSHSGKGIPIHSLGLODPIIARMSICSE--DKKSPSE 1599  
 Db 1712 T-PEEVSSTPAEERK---LYLOTPTSSEKRG-----SPLIOPEPESSEHRESSPRK 1758

QY 1600 CSLIASSPEENMP-----ACOKAYNLNR-TPSVTVLNNNSAPARANQNFDEMEG 1648  
 Db 1759 TSLVYESADNDEPETEERLDEDAAFKGDMPFIEPIYV-----EEFY 1802  
 QY 1649 IRE-----TSOYLIRPSSS-----PPTTIONENLKSMTHKRSQSS 1685  
 Db 1803 IDEHGHVYVKVTKIRIRYVSEGEKEEIMVQMPQEPVNIIEGDDGVYIKRVLKS 1862  
 QY 1686 YTRLKSDPPE 1695  
 Db 1863 DTEQSEBDNNE 1872

Search completed: July 1, 2003, 14:37:02  
 Job time : 61.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 89 seconds  
(without alignments)  
3970.458 Million cell updates/sec

Title: US-10-021-571-2

Perfect score: 8853

Sequence: 1 MSVLISQSVINVEENIPALKEKCDVDERNECGTPTLMLAEGNVEIVKLN 1715

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.podent:\*
- 12: sp.virus:\*
- 13: sp.vertibrate:\*
- 14: sp.unclassified:\*
- 15: sp.rvivirus:\*
- 16: sp.bacteriap:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	8853	100.0	1715	11	09ERD4	09ERD4 rattus norv
2	8639.5	97.6	1762	11	09ERG6	09ERG6 rattus norv
3	8212	92.8	1777	4	09UDH0	09UDH0 homo sapien
4	5073.5	57.3	1031	4	09UDF42	09UDF42 homo sapien
5	2408	27.2	543	4	09H9E4	09H9E4 homo sapien
6	2376.5	26.8	1498	5	09W211	09W211 drosophila
7	1848.5	20.9	1398	5	020109	020109 caenorhabdi
8	1024	11.6	246	4	09NT37	09NT37 homo sapien
9	599.5	6.8	129	4	09H889	09H889 homo sapien
10	555	6.3	426	16	08YTG9	08YTG9 anabaena sp
11	555	6.3	1856	4	099407	099407 homo sapien
12	551	6.2	1943	11	061307	061307 mus musculi
13	545	6.2	1719	4	013768	013768 homo sapien
14	529	6.0	2622	11	070511	070511 rattus norv
15	526	5.9	4377	4	012955	012955 homo sapien
16	516.5	5.8	1848	11	061302	061302 mus musculi

17	510	5.8	1726	11	08VC68	08VC68 mus musculi
18	505.5	5.7	1136	6	09N180	09N180 bos taurus
19	504	5.7	2443	5	09VSA2	09VSA2 drosophila
20	503.5	5.7	843	11	F97582	F97582 rattus norv
21	500.5	5.7	1486	4	08TEF1	08TEF1 homo sapien
22	499.5	5.6	1159	5	09NCP8	09NCP8 drosophila
23	493.5	5.6	833	4	096186	096186 homo sapien
24	492.5	5.6	784	5	09VCA7	09VCA7 drosophila
25	492.5	5.6	4001	5	08WR07	08WR07 drosophila
26	490.5	5.5	6994	5	017343	017343 caenorhabdi
27	488.5	5.5	1009	5	08SWY2	08SWY2 drosophila
28	486.5	5.5	1599	11	09NHO	09NHO mus musculi
29	483.5	5.5	1762	11	088521	088521 rattus norv
30	479.5	5.4	1188	4	09H288	09H288 homo sapien
31	479	5.4	2039	5	017489	017489 caenorhabdi
32	471	5.3	1282	5	08T4F8	08T4F8 drosophila
33	471	5.3	2119	5	09VAU5	09VAU5 drosophila
34	471	5.3	2119	5	08T9G9	08T9G9 drosophila
35	467.5	5.3	1549	5	09V4B1	09V4B1 drosophila
36	464.5	5.2	1549	5	024241	024241 drosophila
37	459	5.2	743	4	08TB46	08TB46 homo sapien
38	455	5.1	460	4	09H6J9	09H6J9 homo sapien
39	454	5.1	1867	5	017486	017486 caenorhabdi
40	453	5.1	627	4	096G77	096G77 homo sapien
41	452.5	5.1	1809	5	017487	017487 caenorhabdi
42	451	5.1	1349	11	08VDN6	08VDN6 mus musculi
43	449.5	5.1	1088	4	013484	013484 homo sapien
44	443.5	5.0	1815	5	017488	017488 caenorhabdi
45	441	5.0	435	4	09HA95	09HA95 homo sapien

## ALIGNMENTS

### RESULT 1

ID 09ERD4 PRELIMINARY: PRT: 1715 AA.

AC 09ERD4: 01-MAR-2001 (TREMREL. 16, Created)

DT 01-MAR-2001 (TREMREL. 16, Last sequence update)

DT 01-JUN-2002 (TREMREL. 21, Last annotation update)

DE Ankyrin repeat-rich membrane-spanning protein.

GN ARMS.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RX MEDLINE=20585245; PubMed=1150334;

RT Kong H., Boulter J., Weber J.L., Lai C., Chao M.V.;

RT "An evolutionarily conserved transmembrane protein that is a novel

RT downstream target of neurotrophin and ephrin receptors.";

RL J. Neurosci. 21:176-185(2001).

DR EMBL: AF313464; AAC34167.1; .

DR HSSP: P80144; ZMYO.

DR InterPro: IPR002110; ANK.

DR Pfam: PF00023; ANK\_12.

DR PRINTS: PR01415; ANKYRIN.

DR SMART: SM00248; ANK\_11.

DR PROSITE: PS50088; ANK\_REPEAT; 10.

DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.

KW ANK repeat; Repeat.

KW SEQUENCE 1715 AA; 190532 MW; 505593CA19A6CDC CRC64;

QY Query Match 100.0%; Score 8853; DB 11; Length 1715;

DB Best Local Similarity 100.0%; Pred. No. 0;

Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSVLISQSVINVEENIPALKEKCDVDERNECGTPTLMLAEGNVEIVKLN 60

1 MSVLISQSVINVEENIPALKEKCDVDERNECGTPTLMLAEGNVEIVKLN 60

```

OY 61 GANCNLEDDNDNTALISASKEGHIHVEELKSGASLEHRMGWGTALMAACVGRPDV 120
    |||||
Db 61 GANCNLEDDNDNTALISASKEGHIHVEELKSGASLEHRMGWGTALMAACVGRPDV 120
OY 121 ELLLSHGANPSVTGLQSVYPIIMAGRGHADIYELLONGAKVCSDKYGTTPVMAAR 180
    |||||
Db 121 ELLLSHGANPSVTGLQSVYPIIMAGRGHADIYELLONGAKVCSDKYGTTPVMAAR 180
OY 181 KGHLECVHLLAMGADVQEGANSMTALIVAKGCTOSVEELLKRNPNVLLTDKGNFA 240
    |||||
Db 181 KGHLECVHLLAMGADVQEGANSMTALIVAKGCTOSVEELLKRNPNVLLTDKGNFA 240
OY 241 LMIAKEGHIIEIVDLNAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
    |||||
Db 241 LMIAKEGHIIEIVDLNAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
OY 301 QDNKTAALYAWVEKGNATVRDILQCNPDTEICTKDGETPLIKATMRNIEVVELLDKGA 360
    |||||
Db 301 QDNKTAALYAWVEKGNATVRDILQCNPDTEICTKDGETPLIKATMRNIEVVELLDKGA 360
OY 361 KVSAYDKKGPPLHVAIRGRSRRLAELLRMPKOSRLLYRPNKAGEPNYNDICSHQSL 420
    |||||
Db 361 KVSAYDKKGPPLHVAIRGRSRRLAELLRMPKOSRLLYRPNKAGEPNYNDICSHQSL 420
OY 421 TQIFGARHLSPETDGDMLGYDLYSSALADILSEPTMQPICYGLYAOVSGSKFLLKL 480
    |||||
Db 421 TQIFGARHLSPETDGDMLGYDLYSSALADILSEPTMQPICYGLYAOVSGSKFLLKL 480
OY 481 EDEMTEFAGQOETEPFQFSMLIVPLTLLCGGLVFPVVDNLAIAISLFTALYIF 540
    |||||
Db 481 EDEMTEFAGQOETEPFQFSMLIVPLTLLCGGLVFPVVDNLAIAISLFTALYIF 540
OY 541 FVIFYGGRGEGSESNMAMALSTRARHIGYLELLEFKLMPNPPELPQTKALPVPLF 600
    |||||
Db 541 FVIFYGGRGEGSESNMAMALSTRARHIGYLELLEFKLMPNPPELPQTKALPVPLF 600
OY 601 TDYNRLSSVGETSLAEMIAATLSDACEREFGLATRLFVRPTEESOGKKKKTKCLPS 660
    |||||
Db 601 TDYNRLSSVGETSLAEMIAATLSDACEREFGLATRLFVRPTEESOGKKKKTKCLPS 660
OY 661 FVIFLFIYGCIIAGITLLAIFRVDPKHLTVNAIILISASVGLAFVLCNCRIMVYDLSL 720
    |||||
Db 661 FVIFLFIYGCIIAGITLLAIFRVDPKHLTVNAIILISASVGLAFVLCNCRIMVYDLSL 720
OY 721 NSORRRLHSAASKLHLKLESEGFMVYLCEVELMARMAKTIDSFQONOTRLVYIIDGDAC 780
    |||||
Db 721 NSORRRLHSAASKLHLKLESEGFMVYLCEVELMARMAKTIDSFQONOTRLVYIIDGDAC 780
OY 781 EODKVLQMLDTRVYLFSGKPPILAFASDPHIIKAINONLNSVLSDSNINCHDMYRNVH 840
    |||||
Db 781 EODKVLQMLDTRVYLFSGKPPILAFASDPHIIKAINONLNSVLSDSNINCHDMYRNVH 840
OY 841 LPVFLNSRGLSNARKFLVTATNGDITCSDTGTQEDTDRRVSONSIGEMTKLGSKALN 900
    |||||
Db 841 LPVFLNSRGLSNARKFLVTATNGDITCSDTGTQEDTDRRVSONSIGEMTKLGSKALN 900
OY 901 RRDYRRRQOMORTITRQMSFDLTKLVTEDMFSDISPOJMRRLINITSYVGRLLRANQIT 960
    |||||
Db 901 RRDYRRRQOMORTITRQMSFDLTKLVTEDMFSDISPOJMRRLINITSYVGRLLRANQIT 960
OY 961 FNNRRLASWIMLTQOWPRTSWMLLYLETEEGLPDQMTLKTWERISKNIPTTKDVPPL 1020
    |||||
Db 961 FNNRRLASWIMLTQOWPRTSWMLLYLETEEGLPDQMTLKTWERISKNIPTTKDVPPL 1020
OY 1021 EIDDDIRNFVEVLESSRTPVLYARDVKTFLPCTVNLDPKLEIILADVAARAEQINIGLAY 1080
    |||||
Db 1021 EIDDDIRNFVEVLESSRTPVLYARDVKTFLPCTVNLDPKLEIILADVAARAEQINIGLAY 1080
OY 1081 PPLPLHGGPPRPSPGSGOPASVCSASNGPFPBGVVSPOPHSSYSGLSGPOHPFNRA 1140
    |||||
Db 1081 PPLPLHGGPPRPSPGSGOPASVCSASNGPFPBGVVSPOPHSSYSGLSGPOHPFNRA 1140

```

```

OY 1141 AVPATSSLLSSMTVDVYCEKRLQTEGLDONMMPOYCTTIKANINGRVLQCNTEDEK 1200
    |||||
Db 1141 AVPATSSLLSSMTVDVYCEKRLQTEGLDONMMPOYCTTIKANINGRVLQCNTEDEK 1200
OY 1201 KEMANNGDWHLEFRSWLLEMRSVESQVVPEDPFLLENSSAPVPHGESARRSRHTLEPLT 1260
    |||||
Db 1201 KEMANNGDWHLEFRSWLLEMRSVESQVVPEDPFLLENSSAPVPHGESARRSRHTLEPLT 1260
OY 1261 ELSSQTPYTLNFSPEBELNTLGLDEGAPRHSNLSWSQTRTPSLSLNSQDSISIEISKLT 1320
    |||||
Db 1261 ELSSQTPYTLNFSPEBELNTLGLDEGAPRHSNLSWSQTRTPSLSLNSQDSISIEISKLT 1320
OY 1321 DKVOAERADYREYIAQMSQLEGTSSTISGRSSPHSTYYIGQSSSGSISTTLEQERG 1380
    |||||
Db 1321 DKVOAERADYREYIAQMSQLEGTSSTISGRSSPHSTYYIGQSSSGSISTTLEQERG 1380
OY 1381 KEGELKQEDGRKSFLLMRGVDYIDSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
    |||||
Db 1381 KEGELKQEDGRKSFLLMRGVDYIDSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
OY 1441 SERPSLEFONDILKLGGLRQKLPSPDEDSGTRQVITPHCSKMITRKLAKKQRECAP 1500
    |||||
Db 1441 SERPSLEFONDILKLGGLRQKLPSPDEDSGTRQVITPHCSKMITRKLAKKQRECAP 1500
OY 1501 QESHAEPITRTFIKAKELYLDALDKRSDSGVRSNESSPNHSLHNEADDSQLEKANLI 1560
    |||||
Db 1501 QESHAEPITRTFIKAKELYLDALDKRSDSGVRSNESSPNHSLHNEADDSQLEKANLI 1560
OY 1561 ELEDGHSGRKMPHSLSGLODPIIARMSICSEDKKSPESCSLIASSPESWPACOKAYN 1620
    |||||
Db 1561 ELEDGHSGRKMPHSLSGLODPIIARMSICSEDKKSPESCSLIASSPESWPACOKAYN 1620
OY 1621 LNRPSTVTNNATPTNANONFDEIGIRETSOYILRPGSPNPTAYONNLKSMHX 1680
    |||||
Db 1621 LNRPSTVTNNATPTNANONFDEIGIRETSOYILRPGSPNPTAYONNLKSMHX 1680
OY 1681 RSQRSSYTRLXSKDASELHAASSESTGFGERESIL 1715
    |||||
Db 1681 RSQRSSYTRLXSKDASELHAASSESTGFGERESIL 1715

RESULT 2
OY 09EOG6 PRELIMINARY; PRT; 1762 AA.
ID 09EOG6
AC 09EOG6;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-OCT-2001 (T-EMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE KIDINS220.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NBL_TaxID-101116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-20568256; PubMed-10998417;
RA Iglesias T., Cabrera-Poch N., Mitchell M.P., Naven T.J., Rozenfurt E.,
RA Schiavo G.;
RT "Identification and cloning of Kidins220, a novel neuronal substrate
of protein kinase D."
RL J. Biol. Chem. 275:40048-40056(2000).
RN [2]
RN SEQUENCE FROM N.A.
RA Schiavo G.;
RN Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Mitchell M., Schiavo G.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF230045; ANG35185.2;
DR HSSP; P80144; 2MYO.
DR InterPro; IPR002110; ANK.

```

DR pfam: PF00023; ank: 12.  
 DR PRINTS: PRO1415; ANKYRIN.  
 DR SMART: SM00248; ANK: 11.  
 DR PROSITE: PS0088; ANK\_REPEAT; 10.  
 DR ANK repeat; Repeat.  
 SK SOURCE 1762 AA; 195715 MW; OCB2689A57JF8AE4 CRC64;

Query Match 97.6%; Score 8639.5; DB 11; Length 1762;  
 Best Local Similarity 95.8%; Pred. No. 0;  
 Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;

QY 1 MSVLIQSIVNYVEENIPALKALLKCKDVERNECGOTPLMLAEOGNVEIVKELKN 60  
 DB 1 MSVLIQSIVNYVEENIPALKALLKCKDVERNECGOTPLMLAEOGNVEIVKELKN 60  
 QY 61 GANCNLEDDNNTALISASKEGHIHYEELKSGASLEHRDNGGWTALMACYKGTDTYV 120  
 DB 61 GANCNLEDDNNTALISASKEGHIHYEELKSGASLEHRDNGGWTALMACYKGTDTYV 120  
 QY 121 ELLSHGAMPSTYGLQYVPIYIWAAGRHADIVHLLONGAKVNCSDKYGTTPYMAAR 180  
 DB 121 ELLSHGAMPSTYGLQYVPIYIWAAGRHADIVHLLONGAKVNCSDKYGTTPYMAAR 180  
 QY 121 ELLSHGAMPSTYGLQYVPIYIWAAGRHADIVHLLONGAKVNCSDKYGTTPYMAAR 179  
 DB 121 ELLSHGAMPSTYGLQYVPIYIWAAGRHADIVHLLONGAKVNCSDKYGTTPYMAAR 179  
 QY 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTOSVKEILKRNPNVNLTDKGN 240  
 DB 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTOSVKEILKRNPNVNLTDKGN 239  
 QY 241 LMAISKEGHEIYODLLDAGTYNIPDRSGDTVLIGAVRGHVEIYRALLQKADIDING 300  
 DB 241 LMAISKEGHEIYODLLDAGTYNIPDRSGDTVLIGAVRGHVEIYRALLQKADIDING 299  
 QY 301 QDNKTLAYMAVEKGNATWADILQCNPDTEICTKDETPLIKATKRNIEVEELLDKGA 360  
 DB 301 QDNKTLAYMAVEKGNATWADILQCNPDTEICTKDETPLIKATKRNIEVEELLDKGA 359  
 QY 361 KYSADVKKGDTPLHVAIRGRSRLAELLRNPKDGLLYRNKAGETPYNIDCSHOKSTL 420  
 DB 361 KYSADVKKGDTPLHVAIRGRSRLAELLRNPKDGLLYRNKAGETPYNIDCSHOKSTL 419  
 QY 421 TOIFGARHLSPTTDDMDGYDLYSSALADILSEPMOPICVGLAOMSGSFTLKL 480  
 DB 421 TOIFGARHLSPTTDDMDGYDLYSSALADILSEPMOPICVGLAOMSGSFTLKL 479  
 QY 481 EDEMKTFAGQOTEPLEFQSWLIVFTLLCGGLGFAFVPVDNTLAIALISFLALITYF 540  
 DB 481 EDEMKTFAGQOTEPLEFQSWLIVFTLLCGGLGFAFVPVDNTLAIALISFLALITYF 539  
 QY 541 FIYIYGGRRGESSMAMALSTRLARHIGYLELFLKMFVNPPELLPEOTYKALPVRLF 600  
 DB 541 FIYIYGGRRGESSMAMALSTRLARHIGYLELFLKMFVNPPELLPEOTYKALPVRLF 599  
 QY 601 TDNRLSSVGGETSLAEMTATLSDACEREGFLATLFRFRREESOGKKKKKCTCLTS 660  
 DB 601 TDNRLSSVGGETSLAEMTATLSDACEREGFLATLFRFRREESOGKKKKKCTCLTS 659  
 QY 661 FVLEFIVGCIIGITLLAIFRVPKHLVYNALISIASVYGLAFNLRTMWOVDSIL 720  
 DB 661 FVLEFIVGCIIGITLLAIFRVPKHLVYNALISIASVYGLAFNLRTMWOVDSIL 719  
 QY 721 NSQRRLHSAASKLHLKSEGFVKCEVELMARNAKTIDSTQNOTRLVVIIDGLAC 780  
 DB 721 NSQRRLHSAASKLHLKSEGFVKCEVELMARNAKTIDSTQNOTRLVVIIDGLAC 779  
 QY 781 EODKVLQMDLTVVLKSGPFIFASDPHIIKAIKONUNSVLRSGNSNGHYMNIYH 840  
 DB 781 EODKVLQMDLTVVLKSGPFIFASDPHIIKAIKONUNSVLRSGNSNGHYMNIYH 839  
 QY 841 LPEFLNSRGLSNARKFLVTSATNGDITCSPTTGTQEDTDRYSONSLGEXTKLGSKTALN 900  
 DB 841 LPEFLNSRGLSNARKFLVTSATNGDITCSPTTGTQEDTDRYSONSLGEXTKLGSKTALN 899  
 QY 901 RROTYYRRQORITTRQMSFDLTKLVTEDEWSDISPQYMRRLNLIVSYTGRLRANOIT 960  
 DB 901 RROTYYRRQORITTRQMSFDLTKLVTEDEWSDISPQYMRRLNLIVSYTGRLRANOIT 959

DB 900 RRDYYRRQORITTRQMSFDLTKLVTEDEWSDISPQYMRRLNLIVSYTGRLRANOIT 959  
 QY 961 FNMDBLAWINTEDEWPRYRTSMILLYLEETGGLDQMTLKMRYERISNIPPTKDVPELL 1020  
 DB 960 FNMDBLAWINTEDEWPRYRTSMILLYLEETGGLDQMTLKMRYERISNIPPTKDVPELL 1019  
 QY 1021 EIDGIRNFEVFLSRTPEVLVARDVKTFLPCTVNDLPRLREIADVRAARQINIGLAY 1080  
 DB 1020 EIDGIRNFEVFLSRTPEVLVARDVKTFLPCTVNDLPRLREIADVRAARQINIGLAY 1079  
 QY 1081 PPLPLEGPPRPPSGYSOPASVCSASFNGPPGGVSPQPHSSYSGLSGPPHFN 1139  
 DB 1080 PPLPLEGPPRPPSGYSOPASVCSASFNGPPGGVSPQPHSSYSGLSGPPHFN 1139  
 QY 1140 -----AAPTATGSSILL 1151  
 DB 1140 FFAPLYPRRYPGGSQHLISRSVYKSLPRDQNGNLPDSDGFMKQROAAVAPATGSSILL 1199  
 QY 1152 SSMVDVYCEKRLQTEGLDQNMMPQYCTTIKKANINGVLSQCNIDELKEMANFGDMH 1211  
 DB 1200 SSMVDVYCEKRLQTEGLDQNMMPQYCTTIKKANINGVLSQCNIDELKEMANFGDMH 1259  
 QY 1212 LFRSVLEMRVYESOVVEDPFLNENSSADVPHEGARSRHTELPLELSSOTPYTLN 1271  
 DB 1260 LFRSVLEMRVYESOVVEDPFLNENSSADVPHEGARSRHTELPLELSSOTPYTLN 1319  
 QY 1272 FSEELNLTGLDEGAPRHSNLSWOSOTRRTSLSLNSODSIEFSLTKDVOAEYRAY 1331  
 DB 1320 FSEELNLTGLDEGAPRHSNLSWOSOTRRTSLSLNSODSIEFSLTKDVOAEYRAY 1379  
 QY 1332 REYIAQMSOLEGTSSTISGRSPHSTYIYGOSSSGSGSIHSTLEOERKGEELKOEDGR 1391  
 DB 1380 REYIAQMSOLEGTSSTISGRSPHSTYIYGOSSSGSGSIHSTLEOERKGEELKOEDGR 1439  
 QY 1392 KSLFLKRGVDYIDSSGYSTNPAFLDITTEDEKSDOSGSKLPGKSSRPSLFTDL 1451  
 DB 1440 KSLFLKRGVDYIDSSGYSTNPAFLDITTEDEKSDOSGSKLPGKSSRPSLFTDL 1499  
 QY 1452 KLGKGLRQKLPSPEDSGTGRVQITPHGSKMITTKRLAK-ORECASPOEHSAPERT 1510  
 DB 1500 KLGKGLRQKLPSPEDSGTGRVQITPHGSKMITTKRLAK-ORECASPOEHSAPERT 1557  
 QY 1511 FIKAREYLSDALDKDSDSGVRSNESPNSHLNEAADSQLEKANILELEDEGHSGK 1570  
 DB 1558 FIKAREYLSDALDKDSDSGVRSNESPNSHLNEAADSQLEKANILELEDEGHSGK 1617  
 QY 1571 RGMPSLSGLQDPIIARMSICSEDKSPSECSLIASPEESWPACQKAYNLNRPSTYTL 1630  
 DB 1618 RGMPSLSGLQDPIIARMSICSEDKSPSECSLIASPEESWPACQKAYNLNRPSTYTL 1677  
 QY 1631 NNNTAPTRANONPEIGIRETSOVLIRPGSPAPRTAVONENLKSMAHRSOSSYTRL 1690  
 DB 1678 NNNTAPTRANONPEIGIRETSOVLIRPGSPAPRTAVONENLKSMAHRSOSSYTRL 1737  
 QY 1691 SKDASELHAASSESTGFGEERSTL 1715  
 DB 1738 SKDASELHAASSESTGFGEERSTL 1762

RESULT 3  
 Q9ULH0 PRELIMINARY; PRT: 1777 AA.  
 AC Q9ULH0;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE KIAA1250 protein (Fragment).  
 GN KIAA1250.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]



RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE=20039619; PubMed=10574462;  
 RA Nagase T., Ishikawa K., Kikuno R., Hirosewa M., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XV.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 6:337-345(1999).  
 DR EMBL: AB033076; BAA86564.2; -  
 DR Interpro: IPR002110; ANK.  
 DR Pfam: PF00023; ank, 12.  
 DR PRINTS: PRO1415; ANKTRIN.  
 DR SMART: SM00248; ANK, 11.  
 DR PROSITE: PS00088; ANK\_REPEAT, 10.  
 DR PROSITE: PS02097; ANK\_REGION, 1.  
 DR ANK repeat; Repeat.  
 KW NON\_TER  
 FT  
 SQ SEQUENCE 1777 AA; 197209 MW; B6505923FB45F143 CRC64;

Query Match 92.8%; Score 8212; DB 4; Length 1777;

Best Local Similarity 89.7%; Pred. No. 0;  
 Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;

DB 1 MSVLISQSYINVEENIPALKALIECKDVERNECGOTPLMAEAGNVEIYKELIKN 60  
 QY 61 GANCLLEPDMNTALISASKEGHIVIEELIKSGASLEHRMGCTALMAAGCYGRDYY 120  
 DB 67 GANCLLEPDMNTALISASKEGHIVIEELIKSGVNEHRMGCTALMAAGCYGRDYY 126  
 QY 121 ELLSHGANPSVTGLQYVYPTIWAAGRHADIVHLLQNGAKVNSDKYGTTPYMAAR 180  
 DB 127 ELLSHGANPSVTGL-YVYPTIWAAGRHADIVHLLQNGAKVNSDKYGTTPYMAAR 185  
 QY 181 KGHLECYVHLLAMGADVQEGANSKTALIVAKGGTOSYKELIKRNPVNLTDKGNTA 240  
 DB 186 KGHLECYVHLLAMGADVQEGANSMTALIVAKGGTOSYKELIKRNPVNLTDKGNTA 245  
 QY 241 LMAISKEGHIETVDLDAGTYVNIPIDRSGPTVLIGAVRGHVEIVRALLOKYADIDRG 300  
 DB 246 LMAISKEGHIETVDLDAGTYVNIPIDRSGPTVLIGAVRGHVEIVRALLOKYADIDRG 305  
 QY 301 QDNKTAIYAWVEKGNATWVRDILQCPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 360  
 DB 306 QDNKTAIYAWVEKGNATWVRDILQCPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 365  
 QY 361 KVSANDKGDPTPLHAIIGRSRRLAELLRNPKGRLLYRNKAGETPTYNIDCSHOSIL 420  
 DB 366 KVSANDKGDPTPLHAIIGRSRRLAELLRNPKGRLLYRNKAGETPTYNIDCSHOSIL 425  
 QY 421 TQIFGARHLSPTETDGMIGYDIYSSALADILSEPTMOPICVGLYKQWGSKGFLLKTL 480  
 DB 426 TQIFGARHLSPTETDGMIGYDIYSSALADILSEPTMOPICVGLYKQWGSKGFLLKTL 485  
 QY 481 EDEKTFPGQOTEPPLPFSMLIVFTLLCGSLGVFAVPVDNLAIALISLSTALIYIF 540  
 DB 486 EDEKTFPGQOTEPPLPFSMLIVFTLLCGSLGVFAVPVDNLAIALISLSTALIYIF 545  
 QY 541 FIVIFGRRGESNNMAMALSTRLARHIGTLELFLKMFVNPELBPOTKALPVRFLE 600  
 DB 546 FIVIFGRRGESNNMAMALSTRLARHIGTLELFLKMFVNPELBPOTKALPVRFLE 605  
 QY 601 TDYNLSVSGETSLAEMIAITLSDACEREFGLATRLFRVFTSESGKKKKTKCCLPS 660  
 DB 606 TDYNLSVSGETSLAEMIAITLSDACEREFGLATRLFRVFTSESGKKKKTKCCLPS 665  
 QY 661 FVIFLFIYGCITAGITLAIIRVDPKHLFVNAIISIASVGLAFVINCRTMWOVDSLL 720  
 DB 666 FVIFLFIYGCITAGITLAIIRVDPKHLFVNAIISIASVGLAFVINCRTMWOVDSLL 725  
 QY 721 NSQRRLHSAASKLHKLSSEGFVKCEVELMARMAKTIDSFQNOTRLVYIIDGLDAC 780

DB 726 NSQRRLHNAASKLHKLSSEGFVKCEVELMARMAKTIDSFQNOTRLVYIIDGLDAC 785  
 QY 781 EOKKVLQMLTVAVLPSKGFIALPASDPHIIIIKAINQNLNSVLRBSNINGDHMYNIYH 840  
 DB 786 EOKKVLQMLTVAVLPSKGFIALPASDPHIIIIKAINQNLNSVLRBSNINGDHMYNIYH 845  
 QY 841 LPVFLNSGRSLNARKFLVTSATNGDITCSPTTGTQEDTQDRVRSQNSLGEKTKGSTALN 900  
 DB 846 LPVFLNSGRSLNARKFLVTSATNGDVPSCPTTGTQEDTQDRVRSQNSLGEKTKGSTALN 905  
 QY 901 RDTYRRRQORPTTROMSFDLTKLVTEDEWFSIDISPTMRRLINIVSYTGILLRANOIT 960  
 DB 906 RDTYRRRQORPTTROMSFDLTKLVTEDEWFSIDISPTMRRLINIVSYTGILLRANOIT 965  
 QY 961 FNMDRLASWNLNLEOMVRYTSWLLVLEFEGIPDQMTLKTMYERISKNIPTTKQVEPL 1020  
 DB 966 FNMDRLASWNLNLEOMVRYTSWLLVLEFEGIPDQMTLKTMYERISKNIPTTKQVEPL 1025  
 QY 1021 EIDGDIRNFEVFLSSRTPVLYARDVTFPLCYVNLDPKLEIITADYRAREQINIGGLAY 1080  
 DB 1026 EIDGDIRNFEVFLSSRTPVLYARDVTFPLCYVNLDPKLEIITADYRAREQINIGGLAY 1085  
 QY 1081 PPLPLEHGPFRPSGYSQPASVCSASFNPGPFGVYSPQPHSYYSGLSGQHPHYN- 1138  
 DB 1086 PPLPLEHGPFRPSGYSQPASVCSASFNPGPFGVYSPQPHSYYSGLSGQHPHYN- 1145  
 QY 1139 -----RAANP 1143  
 DB 1146 FPAAPLYTPRYVGGSOHLISRPVKTSLPDRQONCLEVIKEDAAEGLSSPPDSKSGSP 1205  
 QY 1144 ATGSSLLSSMTVDVVCCEKRLQTEGLDQNMMPQYCTTIKANKINGVLSQCINDELKCKM 1203  
 DB 1206 APGPVVLNLANVDANCKEKLQTEGLDQSLPQYCTTIKANKINGVLSQCINDELKCKM 1265  
 QY 1204 AANFGDWHLFRSVNLEKRSYVSPEDPAPFLNENSAAPVPHGESKRRSSHELPLETSL 1263  
 DB 1266 AANFGDWHLFRSVNLEKRSYVSPEDPAPFLNENSAAPVPHGESKRRSSHELPLETSL 1325  
 QY 1264 SOTPYTLNFSFEELNTGLDEGAPRHSNLSWQOTRRTPLSLNSQDSIEISKLTQDV 1323  
 DB 1326 SOTPYTLNFSFEELNTGLDEGAPRHSNLSWQOTRRTPLSLNSQDSIEISKLTQDV 1385  
 QY 1324 QAEYRDAYREYIAQMSOLEGTSSTISGRSPHSTYIIGSSSGGSIHSTLEOERKXG 1383  
 DB 1386 QAEYRDAYREYIAQMSOLEGTSSTISGRSPHSTYIIGSSSGGSIHSTLEOERKXG 1445  
 QY 1384 ELKQEDGRKSFLLKRGDIVIDSSGVSSTYNDASPLDITDEDEKSDQSGSKLLPGKKSER 1443  
 DB 1446 ELKQEDGRKSFLLKRGDIVIDSSGVSSTYNDASPLDITDEDEKSDQSGSKLLPGKKSER 1505  
 QY 1444 PSLFQTDILKKGGLRYOKLPSEDESGCGRAVOTPHCSKMITRKLKAKORECASPOH 1503  
 DB 1506 PSLFQTDILKKGGLRYOKLPSEDESGCGRAVOTPHCSKMITRKLKAKORECASPOH 1564  
 QY 1504 SAEPITFTIAKEYISDALIDKKDSSDGVRSNENSPHNSHLNEAADSOLEKANLIELE 1563  
 DB 1565 SAEPITFTIAKEYISDALIDKKDSSDGVRSNENSPHNSHLNEAADSOLEKANLIELE 1624  
 QY 1564 DEHSGKRGPHSLSLQDPIIARMSITCEDKSPSEGLIASSPEESPACOKAYNINR 1623  
 DB 1625 DEHSGKRGPHSLSLQDPIIARMSITCEDKSPSEGLIASSPEESPACOKAYNINR 1684  
 QY 1624 TPSTVTLNNNTAPNTANANPDEIEIRTSOVIILRPGPSPTAVONENLSMAHKRSQ 1683  
 DB 1685 TPSTVTLNNNTAPNTANANPDEIEIRTSOVIILRPGPSPTAVONENLSMAHKRSQ 1744  
 QY 1684 RSSYTRLSDASELH-AASESTGGEERESIL 1715  
 DB 1745 RSSYTRLSDASELH-AASESTGGEERESIL 1777

RESULT 4  
 090F42

ID 09UP42 PRELIMINARY; PRT; 1031 AA.  
 AC 09UP42;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Hypothetical 115.3 kDa protein.  
 GN DKFPA34F0621.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL133620; CAB63746.1;  
 DR HSSP; P80144; 2MYO.  
 DR InterPro: IPR002110; ANK.  
 DR Pfam: PF00023; ank; 12.  
 DR PRINTS; PR01415; ANKYRN.  
 DR SMART; SM00248; ANK; 11.  
 DR PROSITE; PS50088; ANK\_REPEAT; 10.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR ANK repeat: Hypothetical protein; Repeat.  
 SO SEQUENCE 1031 AA; 11531 MW; D9E866461C13A2F5 CRC64;

Query Match 57.3%; Score 5073.5; DB 4; Length 1031;  
 Best Local Similarity 95.2%; Pred. No. 0;  
 Matches 974; Conservative 23; Mismatches 13; Indels 13; Gaps 1;

QY 1 MSVLISQSVINVEEENIPALKECKDYDERNECGQPPMLAAGVNEIVKELTKN 60  
 DB 1 MSVLISQSVINVEEENIPALKECKDYDERNECGQPPMLAAGVNEIVKELTKN 60  
 QY 61 GANCNLEDDLNMTALISASKEGHIIVEELLSKSGSLHNDMGMTALMACYKGRDYY 120  
 DB 61 GANCNLEDDLNMTALISASKEGHIIVEELLSKSGSLHNDMGMTALMACYKGRDYY 120  
 QY 121 ELLISHGANSVYGLQSVYPIIWAAGRHADIVHLLONGAVNCSDKYGTPLVWAAR 180  
 DB 121 ELLISHGANSVYGLQSVYPIIWAAGRHADIVHLLONGAVNCSDKYGTPLVWAAR 180  
 QY 181 KGHLECYKHLAMGADVDOGANSMTALIAVKGQYQSVKEILKRPNVNLDDKQNTA 240  
 DB 181 KGHLECYKHLAMGADVDOGANSMTALIAVKGQYQSVKEILKRPNVNLDDKQNTA 240  
 QY 241 LMISKEGHIIVEYDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300  
 DB 241 LMISKEGHIIVEYDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300  
 QY 301 QDNKTALYWAVERKGNATWVDILQCNPDTEICTYKDGSTPLIKATKMNIEVELLDKGA 360  
 DB 301 QDNKTALYWAVERKGNATWVDILQCNPDTEICTYKDGSTPLIKATKMNIEVELLDKGA 360  
 QY 361 KVSADVKKGDTPLVATIRGSRRLAELLRNPKDRLLYRNKGEPPYNDCHOKSIL 420  
 DB 361 KVSADVKKGDTPLVATIRGSRRLAELLRNPKDRLLYRNKGEPPYNDCHOKSIL 420  
 QY 421 TQIFGARHLSPTETDGMGLGYDLSALADILSEPTQOPICVGLYQWQSGSKFLKLKL 480  
 DB 421 TQIFGARHLSPTETDGMGLGYDLSALADILSEPTQOPICVGLYQWQSGSKFLKLKL 480  
 QY 481 EDEKMTFAGQOTEPPLQFQSWLIVLTLGGLGLVAFPPDVTMLAIAISFLALYIF 540  
 DB 481 EDEKMTFAGQOTEPPLQFQSWLIVLTLGGLGLVAFPPDVTMLAIAISFLALYIF 540  
 QY 541 FIVYFGGRGEGESNNAMALSTRLARHIGYLELFLKLMFNPELPEQTKALPYRFLF 600  
 DB 541 FIVYFGGRGEGESNNAMALSTRLARHIGYLELFLKLMFNPELPEQTKALPYRFLF 600  
 QY 601 TDYNNLSVSGETSLAEMIATLSDACEREFGLATRLFRVPRTEESOGKKKKTKCLPS 660  
 DB 601 TDYNNLSVSGETSLAEMIATLSDACEREFGLATRLFRVPRTEESOGKKKKTKCLPS 660

DB 601 TDYNNLSVSGETSLAEMIATLSDACEREFGLATRLFRVPRTEESOGKKKKTKCLPS 660  
 QY 661 FVIFPLFYGCIIAGITTLAIFRVDPKHLTVANAIIISIVYGLAFVNLCTRMQVDSIL 720  
 DB 661 FVIFPLFYGCIIAGITTLAIFRVDPKHLTVANAIIISIVYGLAFVNLCTRMQVDSIL 720  
 QY 721 NSQRRRLNASKRLKLSKSEGMKYLKCEVELMAMAKTIDISFTQNTRLVVIIDGLDAC 780  
 DB 721 NSQRRRLNASKRLKLSKSEGMKYLKCEVELMAMAKTIDISFTQNTRLVVIIDGLDAC 780  
 QY 781 EODKYLQMLDVTYVLESGPPTAIFASDPHIIKAIINONLSVLDNSINGHDYRNIVH 840  
 DB 781 EODKYLQMLDVTYVLESGPPTAIFASDPHIIKAIINONLSVLDNSINGHDYRNIVH 840  
 QY 841 LPVPLNSRLSNARKFVYTSATNGDITGSDPTQOEDMDRVSONSLGEMTKISKALN 900  
 DB 841 LPVPLNSRLSNARKFVYTSATNGDITGSDPTQOEDMDRVSONSLGEMTKISKALN 900  
 QY 901 RRDYRRROMQRTIRONSFDLTLLVTEDMFSDISPTMRRLNIVSVTGRLLRANOIT 960  
 DB 901 RRDYRRROMQRTIRONSFDLTLLVTEDMFSDISPTMRRLNIVSVTGRLLRANOIT 960  
 QY 961 FWMDRASWINTEQMPYRTSWLIIYLETBGLPDQMTLKYER-----IS 1007  
 DB 961 FWMDRASWINTEQMPYRTSWLIIYLETBGLPDQMTLKYERCGADSCDRDRIGIS 1020  
 QY 1008 KNI 1010  
 DB 1021 KSV 1023

## RESULT 5

ID 09H9E4 PRELIMINARY; PRT; 543 AA.  
 AC 09H9E4;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE CDNA FLJ12811 f1s, clone NT2RP2002475.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK022873; BAB14285.1; - F5D643D5A20C641D CRC64;  
 SO SEQUENCE 543 AA; 59976 MW; F5D643D5A20C641D CRC64;

Query Match 27.2%; Score 2408; DB 4; Length 543;  
 Best Local Similarity 86.8%; Pred. No. 1,3e-148;  
 Matches 472; Conservative 24; Mismatches 46; Indels 2; Gaps 2;

QY 1173 MNPQCTTKKANINRGVLSQCNIDELKEMANFEDMHLFMSVTEMSVSQVPPEDP 1232  
 DB 1 MNPQCTTKKANINRGVLSQCNIDELKEMANFEDMHLFMSVTEMSVSQVPPEDP 1232  
 QY 1233 RFLNENSAFVPGESARSSSHLELPTELSTQPTLNFSPFELNTGLDGAAPRSHNL 1292  
 DB 1233 RFLNENSAFVPGESARSSSHLELPTELSTQPTLNFSPFELNTGLDGAAPRSHNL 1292  
 QY 61 RFLSESSGAPRGEFARRARRASHNELHTELSQPTTLNFSPELNTGLDGAAPRSHNL 120  
 DB 61 RFLSESSGAPRGEFARRARRASHNELHTELSQPTTLNFSPELNTGLDGAAPRSHNL 120  
 QY 1293 SMOQSTRPPLSLNSLNSQDSISIKLTKVQAEYDAYREYIAQNSQLEGSTISG 1352  
 DB 1293 SMOQSTRPPLSLNSLNSQDSISIKLTKVQAEYDAYREYIAQNSQLEGSTISG 1352  
 QY 1353 RSSPHSYIYIGQSSSGSISHTEOERKGEKGLKQDGRKSPLMKRGDVIYSSGVSNN 1412  
 DB 1353 RSSPHSYIYIGQSSSGSISHTEOERKGEKGLKQDGRKSPLMKRGDVIYSSGVSNN 1412  
 QY 1412 RSSPHSYIYIGQSSSGSISHTEOERKGEKGLKQDGRKSPLMKRGDVIYSSGVSNN 1412  
 DB 1412 RSSPHSYIYIGQSSSGSISHTEOERKGEKGLKQDGRKSPLMKRGDVIYSSGVSNN 1412

```

QY 1413 EASPLDPTTEDEKSDSGSKLLPGKSSERSTLPOTDLKLGGLRQKLPSPDESGT 1472
DB 241 DASPLDPTTEDEKSDSGSKLLPGKSSERSTLPOTDLKLGGLRQKLPSPDESGT 300
QY 1473 GNVOTPHSCMKIRTKAKORCASPOEHSAPRTFIRAKETSLDALKDKSSDSG 1532
DB 301 EESDMP-LLKDKDKRKAAGKVERVPKSPHSAEIRFIRAKETSLDALKDKSSDSG 359
QY 1533 VASNSPNSHSHNNAADDSOLEKANLLEDEGHSKRGKMPHSLSGLDPTIARMSICS 1592
DB 360 VASNSPNSHSHNNAADDSOLEKANLLEDEGHSKRGKMPHSLSGLDPTIARMSICS 419
QY 1593 EDKSPSECSLIASSPESWPAQCAKAYNLNRPSTVTNNPTANNRANONDEDEGIRE 1652
DB 420 EDKSPSECSLIASSPESWPAQCAKAYNLNRPSTVTNNPTANNRANONDEDEGIRE 479
QY 1653 TSQVILRPGSPNPFAVONENILKSAHRKRSQSSYTRLSKDPASSELH-AASSESTGFBER 1711
DB 480 TSQVILRPGSPNPFAVONENILKSAHRKRSQSSYTRLSKDPASSELH-AASSESTGFBER 539

```

```

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin C.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003453; AAF46710.1; -.
DR HSSP: P42773; 11HB.
DR Flybase: FBgn0034615; CG10074.
DR Interpro: IPR002110; ANK.
DR Interpro: IPR000847; HTH_LysR.
DR Pfam: PF00023; ANK. 10.
DR SMART: SM00248; ANK. 10.
DR PROSITE: PS50088; ANK_REPEAT. 9.
DR PROSITE: PS50297; ANK_REPEAT_REGION. 1.
DR PROSITE: PS00044; HTH_LysR_FAMILY; UNKNOWN_1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1498 AA; 165031 MW; 7C2A888E1B64982C CRC64;

```

Query Match 26.8%; Score 2376.5; DB 5; Length 1498;  
 Best Local Similarity 37.6%; Pred. No. 8.4e-146;  
 Matches 552; Conservative 272; Mismatches 527; Indels 117; Gaps 36;

```

RESULT 6
Q9W211 PRELIMINARY: PRT: 1498 AA.
AC 09W211:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG10074 protein.
GN CG10074.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RX MEDLINE=20196006; PubMed=10731132.
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Doonan K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.C., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinelt K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```

```

QY 1 MSVLISQSVNYVEENIPALKALE-KCKDVERNECCOTPLMAEONGVEIYKELK 59
DB 23 MGSIGRALLQYIDNNNDISGLRAILIDSDRENNATYLVAVAGRLAFREFLA 82
QY 60 NGANCMIEDDNTALISAKREGHIVELKSGSLBHRMGGTATLMAACYGRDY 119
DB 83 RGDVQAEEDDNTALISAKRNLHVDLLDHGAEVHRMGGTATLMAACYGRDY 142
QY 120 VELLISHGANPSVTGLQSYVPIYAGRHADYHLLDQNAKYNSDKYGTPLWAA 179
DB 143 VRLLDKADGNAGH-NYHIGALLMAAGRGYDIYELLVQBARAYVGDYGTALWAC 201
QY 180 RKHLBCVHLLAMGADVQEGANSMTALIVAKGYTOSVEILLKRNPNVLTOKDGT 239
DB 202 RRGVVELVDTLAKAGANVDTAGKYSWPLVYVAAAGHTCVSSILEKRNVALDKDGT 261
QY 240 ALMIASKEGHEIVDLDAGTYVNIPIRSGDTVLIGAVRGHVEIVRALOKYADIDR 299
DB 262 ALCIASREGFODIAASLIAAGAYINIDRGADTPLIHAKAGHRTVEALLKHAIDVIO 321
QY 300 GDNKALYAVAKGNAVVRIOLOCNPTETCTDGEPLIKAKRMRIEVELLDKG 359
DB 322 GKDKRTAITYAVEKHTPTVKKLLATNPDESATKDGDTPLRAVRNRLEIVHLLDRK 381
QY 360 AKVASDKKGDPLVAIRGSRRLAELLRNPKDGRLLRPNKAGETPYNDCSHOKSI 419
DB 382 AKVTASDKKGDPLCIHAMARSKTIVELLRNPKSOLLRANKGEPYNDISLHOTTI 441
QY 420 LTDFGASRLSPETDGMVGYDYSSALADILSEPTMOPPCVGLYQWGSFKLLK 479
DB 442 LGGVFGARLNNNESEGMLGYELYSALADVISEPTLTPTTVGLYAKWGSFKLLK 501
QY 480 LDEKMTFGQGTPELPFQSWLIVETLLTLLCGSLVAVAFV-DINLAIALISLALTY 538
DB 502 LDEKMTFGQGTPELPFQSWLIVETLLTLLCGSLVAVAFV-DINLAIALISLALTY 561
QY 539 IFFIYVPGREGESWMMANALSTRLARHIGLELFRKMFVNPPELPEOTTRALPYRF 598
DB 562 LLLAAYRCNVQ-MDMQWASVQDGLKRMRLRLILQVAFCHPG-PODSQAQAPYRF 618
QY 599 LFTDYNRLSSVGGESLAEMLATLSDACEREGFLATLFRVFRTE-ESOGKKRWKTC 656
DB 619 HFAEANSASPT-GDAVAVHMLAALADATESHYGWLATRLYAFAPKCLKDVGVGRWRMC 677
QY 657 CLPSVFIPEFYICGIIAGITLAI-RRDPKHLVYNNILISISVGLAVLNCRT 711
DB 678 CLPIYLFELALVYVGTISLVAYFTFADEKEKEHILV-ALYVIAVWGTLLICTHNV 735

```





DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)  
 DE Hypothetical protein A112748.  
 GN A112748.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 CX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneke T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,  
 Watanabe A., Itiguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 Yasuda M., Tabata S.;  
 RT \*complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium Anabaena sp. strain PCC 7120.\*;  
 RL DNA Res. 8:205-213(2001).  
 EMBL: AP003590; BAB7447.1; -.  
 InterPro: IPR002110; ANK.  
 Pfam: PF00023; ank. 13.  
 PRINTS: PR01415; ANKYRIN.  
 SMART: SM00248; ANK. 13.  
 DR PROSITE: PS50088; ANK\_REPEAT. 11.  
 DR PROSITE: PS50297; ANK\_REGION. 1.  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 426 AA; 43800 MW; CF089E5D43CAA03 CRC64;

Query Match 6.3%; Score 555; DB 16; Length 426;  
 Best Local Similarity 37.2%; Pred. No. 8.0e-28;  
 Matches 149; Conservative 67; Mismatches 177; Indels 8; Gaps 6;

QY 9 VINVEEENIPALKLEKCKDYDERNECGOTPLMAEGQNEIYKELKNGANCNL-E 67  
 DB 8 LKVAKSGDKKGLGALLAAGVDCIDRGGTTALMPRANGTEIYKSLDGGANVNLAR 67  
 QY 68 LLDNMTALISASKEGHIHVEELKSGSLHEDMGMTALMAWYCKRTDVELLSHG 127  
 DB 68 KRYGLFALMLAASANGVDIYQLLISGAANVANNEDGSTALMAALKNGVEARVLLAAG 127  
 QY 128 ANSVYVGLQSYVPIIWAARGADIVHLLONGAKVNCDKGTTPVLAARKHLECY 187  
 DB 128 ADVNITDKDDT-ALKLAVERGQAAYVQLLPFGADANCBDEGETLLMLAADSGHDV 186  
 QY 188 KHLAMGADYDEGANSMTALIVAKGYTQSYKEIKRPNVNLDDKDGNTALMASKE 247  
 DB 187 QVLLAAGVDVNEQDGTALIAVAAGNAIKIILDRADVNHQDQESLHATVE 246  
 QY 248 GHIEIVDLDAGTYVNIIPRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGOD-NKTA 306  
 DB 247 GYVDVYVGLLNGQANTQIKKRLGDTPLVALQGHQIVETLL-KYG-ANVHGDNLGETP 304  
 QY 307 LYMAVEKGNATMWDILQCPDPEICKDEETPLIKTKRKNIEVVELLDKAKYSAVD 366  
 DB 305 LTLAASGHTAYVAILLDYCANANIPASDKTALIKTEHHGCVIQLLAKGANVYOD 364  
 QY 367 KKGDTPLHVAIRGSRRLAELLRNPKDGLLYRPNKAGET 407  
 DB 365 SVGATALIMASGYNKVVQILLLEGADTNL---KNGGYT 402

RESULT 11  
 Q99407 PRELIMINARY; PRT; 1856 AA.  
 ID 099407;  
 AC 099407;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Ankyrin.  
 GN ANK1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 CX NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97382244; PubMed=9235914;  
 RA Gallagher P.G., Tse W.T., Scarpa A.L., Lux S.E., Forget B.G.;  
 RT "Structure and organization of the human ankyrin-1 gene. Basis for  
 RT complexity of pre-mRNA processing.";  
 RL J. Biol. Chem. 272:19220-19228(1997).  
 DR EMBL: U50133; AAB47805.1; -.  
 DR EMBL: U50093; AAB47805.1; JOINED.  
 DR EMBL: U50094; AAB47805.1; JOINED.  
 DR EMBL: U50095; AAB47805.1; JOINED.  
 DR EMBL: U50096; AAB47805.1; JOINED.  
 DR EMBL: U50097; AAB47805.1; JOINED.  
 DR EMBL: U50098; AAB47805.1; JOINED.  
 DR EMBL: U50099; AAB47805.1; JOINED.  
 DR EMBL: U50100; AAB47805.1; JOINED.  
 DR EMBL: U50101; AAB47805.1; JOINED.  
 DR EMBL: U50102; AAB47805.1; JOINED.  
 DR EMBL: U50103; AAB47805.1; JOINED.  
 DR EMBL: U50104; AAB47805.1; JOINED.  
 DR EMBL: U50105; AAB47805.1; JOINED.  
 DR EMBL: U50106; AAB47805.1; JOINED.  
 DR EMBL: U50107; AAB47805.1; JOINED.  
 DR EMBL: U50108; AAB47805.1; JOINED.  
 DR EMBL: U50109; AAB47805.1; JOINED.  
 DR EMBL: U50110; AAB47805.1; JOINED.  
 DR EMBL: U50111; AAB47805.1; JOINED.  
 DR EMBL: U50112; AAB47805.1; JOINED.  
 DR EMBL: U50113; AAB47805.1; JOINED.  
 DR EMBL: U50114; AAB47805.1; JOINED.  
 DR EMBL: U50115; AAB47805.1; JOINED.  
 DR EMBL: U50116; AAB47805.1; JOINED.  
 DR EMBL: U50117; AAB47805.1; JOINED.  
 DR EMBL: U50118; AAB47805.1; JOINED.  
 DR EMBL: U50119; AAB47805.1; JOINED.  
 DR EMBL: U50120; AAB47805.1; JOINED.  
 DR EMBL: U50121; AAB47805.1; JOINED.  
 DR EMBL: U50122; AAB47805.1; JOINED.  
 DR EMBL: U50123; AAB47805.1; JOINED.  
 DR EMBL: U50124; AAB47805.1; JOINED.  
 DR EMBL: U50125; AAB47805.1; JOINED.  
 DR EMBL: U50126; AAB47805.1; JOINED.  
 DR EMBL: U50127; AAB47805.1; JOINED.  
 DR EMBL: U50128; AAB47805.1; JOINED.  
 DR EMBL: U50129; AAB47805.1; JOINED.  
 DR EMBL: U50130; AAB47805.1; JOINED.  
 DR EMBL: U50131; AAB47805.1; JOINED.  
 DR EMBL: U50132; AAB47805.1; JOINED.  
 DR HSSP: Q00420; IAWC.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000488; Death.  
 DR InterPro: IPR000906; ZUS.  
 DR Pfam: PF00023; ank. 23.  
 DR Pfam: PF00531; ZUS. 1.  
 DR Pfam: PF00791; ZUS. 1.  
 DR PRINTS: PR01415; ANKYRIN.  
 DR SMART: SM00248; ANK. 21.  
 DR SMART: SM00005; DEATH. 1.  
 DR SMART: SM00218; ZUS. 1.  
 DR PROSITE: PS50088; ANK\_REPEAT. 20.  
 DR PROSITE: PS50297; ANK\_REGION. 1.  
 DR PROSITE: PS50017; DEATH\_DOMAIN. 1.  
 KW ANK repeat; Repeat.  
 SQ SEQUENCE 1856 AA; 203466 MW; 4A4E3A794DD6401F CRC64;

Query Match 6.3%; Score 555; DB 4; Length 1856;  
 Best Local Similarity 20.6%; Pred. No. 1.1e-26;  
 Matches 395; Conservative 269; Mismatches 644; Indels 612; Gaps 74;  
 QY 21 LKALLEKCKDYDERNECGOTPLMAEGQNEIYKELKNGANCNLELDNMTALISASK 80  
 DB 61 VRELIVYGANVNAOSQGTPLPYMAAQNHELVKFLLENGANQNVATEDGETPLAVALQ 120



```

81 EGH-----83
121 QGHENVAHLNNGTKGVRLPALHIAARNDFTAAVLONDPDVLSTGFTPLHIA 180
84 -----IHIVEELLSGASLEHRDMGWTALMACYKGRDVELLSHGANSPTGLQYS 138
181 AHENLVNAQOLLNNGASVNFPTPONGITPLHIASRGVIMVRLLDGNAQIE-TKTKDE 239
139 YPIITMAKGRHADIVHLLONGAKVNCSDKYGTPPLVMAARKGHECVKHLAMGADV 198
240 LTPHCAARNGRHVRSEILDHGAPICAKTNGSLPHMAAQGHLCVRLLOYDAEID 299
199 -----ORGA-----NSMTALIVAKGYTQSVKELK 225
300 DITLDHLPPLVAAHCHHRAKAYLLDKAKPNSBALGFTPLHIAKQOVAVARSLLOYG 359
226 RNPVNLDDKQNTALMIASKEGHEIYODLLDAGTVNIPDRSGDPTVLGAVRGHVEI 285
360 TGASIDAVTESGLFPLHVASFMGHLPIYKNLLDQASPNVSNVKEFPLHMAARAGHEV 419
286 VRALLQKADDIRQDNKKTALYNAVEKGNATWVRDILQCPDEICTKQDEPTPLIATK 345
420 AKYLLONKAKANAKAKDQTPHCAARIGHNNMKLENNANENLATTTAGHTPLHIAAR 479
346 MNIEVVELLDKAKYSAVDKGDTPHVAIRGRSRRLAELLRNPDGRLTYRPNKAG 405
480 EGHVETVALLEKEKESQACMTKGTPLHIAKGYKAVALLLE-RDA-----HPNAG 533
406 E---TPYIDCSHO-----KSIL-----TOIFGAR-----427
534 KNGLPPLHVAHNNMLDVLKLLPRGSPHSPANNGYTPLHIAKQOVAVARSLLOYG 593
428 -----HLSPETDMDLYGDLSSALADLSEPTMP-----PIC--462
554 SANASVCGTPLHIAOGBAEMVALLSKOANGNNGKSGPLHIAVOEGHVPADV 653
463 -----VGLYAQWGSK--SFLIKLED-EMKTFAGOOTEPPLQFSWL 501
654 LIKHGVNDATTRMGYTPHVAHSHYGNIKLYKFLQOADVNAATKIGY--SPLHQAQ 711
502 --LYFLTLGGGLVFAFVVD-----TMLAIAISFLALYIFFIYFGGRGE 553
712 GHTDVTLLKNG-----ASPNEVS SDGTPPLAIKRLGYISVDVLYV-----756
554 SMNAMALSTRLARHIGYELLFKLMFVNPEPEQTKALPVFLPTDYNRLS-SYGE 612
757 TDESFVLYSDKHR--MSPEYVDELIDVSEDEGTAHITIMGEELISKRAERDSRDVEE 815
613 TSLAEMITATLSDACEREGFLATRLFR-----VETEE--SQGKKMKKTCCLPSFV 662
816 KELIDFVPLDOVE-----SPAIPRIKAMPETVIRSEEGQAKKEDEBLSIPS--867
663 IFLFVCGITINGITLLAIFRVDPKHLVNALISIASVGLAFVLCNRTMQUVIDSLNS 722
868 -----SPATEISDNI--SPVASPVHGFVVSF-----WYDAGS 900
723 QKRRLHSAASKLHKLSRGMKVLKCEVELAKARAKTIDSTQOTRLVYIIDLADCEQ 782
901 MRGRHN-----GLRVIIP-----PRICAAFTKTKLY-----929
783 DKVLQMLDVTAVLESKGPPIAIFASDPHIIKAINOMINSLVLDNSNIGHDYMN--IVHL 841
930 --KPKLSTPPL-----AEEGGLASRIIALGPTGAQPLSPYIVEI 968
842 PVFL-NSRGLSNAKFLVTSATNDITCSPTGTQOEDTDRVYSONSIGEMTKLGSKTALN 900
969 PHFASHG--DRELVVLRSENGSV-----WKHRSRYGHSYLDQILN--GMDELG 1016
901 RRDYRRORQRTTROMSFDLTKLVT-----EDWFSISPOTMRLNLIYSVGRLLRAN 957
1017 SLELEKRRKCRIT--TDPLLYIVMSRLQD-YDITGEG-----GSLSKVLPLY 1066

```

```

958 QITPNMDRLASWINTLEQMPYRTSWLILYLETGEGPDQMTLK-----TMYER 1005
1067 QATFENAVATKRLAQ-----AQVPDELVTKLGNQATFSPYVTEPR 1112
1006 ISK-----NIPTKDVEPLEIDGDIRNEFEV-----SSRTPLVVA 1042
1113 RKFHNRIGIRIPLPSWMDNPNRDSGEGTTLRLCSYIGGTDQAWEDITGKLYVA 1172
1043 RDVKTPLPCTVINIDPKRETIADVRAARQINIGLAYPPLHIEGPPRPGYOPASV 1102
1173 NECANF---TTNV--SARFMLSOPRTAAVAVETALLKEL-----TAVPYMAKF 1217
1103 CSSASFNGPVGVSPPQHSYSGLSGPQHPFYRAAVPARGSSLLSTVTVYCEK 1162
1218 VIFAKMNDREG-----RLRCYKMTDDKVDYT 1244
1163 LRO-----IEG-----LDONMPOCYTTRKAN-----INGRYL 1191
1245 LEQHENFVEVARSDIEVLEGMSLFAELSGNLVP-----YKKAQOQSRFHQSPRENRLA 1299
1192 SOCNIDELKE-----MANNFGD--WHLRSMVLEKRSVESOVVPEPR-----1233
1300 MPYKVRDSSREPGSLSPFLRKAMKYEDTOHILCHLNTMPCAKSGAGADRRTPTPLAL 1359
1234 ---FLNENSGAPVPHEGSAAR-----SHTLPLTFLSQTPTLNFSPFELNTGLD--1283
1360 RYSLIESPTGSLSGTEQAEKMAVISEHLGLSMELARE---LOFSYEDINRLRVENP 1415
1284 ---EAPRHSNLSWOSQTRRPPSSLNSODSSEIFSKLTKVOA-----EYRD 1329
1416 NSLLEQSVALLNL--WYIRGONAMNENLTALOSIDREIVNMLDESGRSHNLKPRDRH 1474
1330 AVEEYIAQSOLEGGGSSISGRSSPHSTYIYGSSGSGSHHTLEDRGREGELKQED 1389
1475 TDRDYLSPSQMNG--YSSLODELSP-----ASLCAALSPRLADQ-----1514
1390 GRKSFMLKRGVDYDSSGVTNEASPLDPI-----TEED--EKSPDS--GSKLPGKKS 1440
1515 -----YNNVAVADAIPLAATHEHTMLEMSDMQVMSGLRPLSYT 1554
1441 SERPSLFQTDLIKGGGLKAYOKLPSDEDSGGRVOITPHCSKMITRKLAKQRECAP 1500
1555 AEDSSD-----KCSKAEDSDATGH-----EMKLEGLALSEEPGP 1588
1501 QEHSAPRITPFIKAKETLSDA---LIDKSDSSDGSVRNESSPNHSLHNEADDSOLEKA 1557
1589 ELGSLD---LYEDDTVDSDATNGTDLLE--QEBGORSSEKLPGRKODDAGACQ--DSE 1642
1558 NLELEDEHSGSKRGMPHSLSGLOPPIIARMSIGSEBK-----KSPSECSLIASSPERSW 1612
1643 NEVSLVSGHQRGQARTHS-----PTVQVTERSGDRQLQMDADAGSIYSYLDAAQCSW 1696

```

RESULT 12

061307 PRELIMINARY; PRT; 1943 AA.

AC 061307; 061305; 061306; 061308; 061309; 061310; 008866; 008867;

DT 01-JUN-1998 (TREMblrel. 06, Created)

DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)

DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DN Ankyrin 3 (Ankyrin 3) (Epithelial ankyrin) (Ankyrin-3).

GN ANK3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1-6).

RC STRAIN=C57BL/6J; TISSUE=KIDNEY;

RX MEDLINE=95340633; PubMed=7615634;

RA Peters L.L., John K.M., Lu F.M., Eicher E.M., Higgins A., Yialamas M.,

RA Turtzo L.C., Otsuka A.J., Lux S.E.;

RT "Ank3 (epithelial ankyrin), a widely distributed new member of the





```

Db 724 GNKIYVFLQSAKVNKTKNGYTALHQAQGHTHIINVLQNNASPNELTVNG----- 779
QY 517 FAFPDVNLAIISFLALITYFI----- 544
Db 780 -----NTALAIARRGIIYVDTLKVTEEIWTITTEKHMANVPETMNEVLDMDEY 834
QY 545 -----YFGREGESNMAMALSTRLARHIGYLELFLKMFVNPPELPEOT 591
Db 835 KASAPKLSDEYISDGEED-----ALIGDTDKYLIG-----PODLKEGD 877
QY 592 KALPVAFLETTYNRLSSVGETSLAEMIATLSDACEREFGLATRLFAVFRTEESOGKK 651
Db 878 DLPAPR-----GYG--FSLGARASLR-----SSSDRSTYLNASSAR----- 915
QY 652 WKTCTCLSFVFLFYVCGIAGTLLAIFRVP-KHLT-----VANILISIASVYG 702
Db 916 -----DSMIEELIPLVPSKEQHLTFTRFEDSDSLRHYSMADTLDNVNLVSPVHSGL 968
QY 703 LAFVLCRPMQVYDLSLNSQ-----RK-----RLMSASKLHL-----KSE 740
Db 969 VSFVADAKG-----GSMGSRHGMRIIPRCKTAPTRICRLVKRKLANPPMVEGE 1023
QY 741 G-----FMKVLKCEVELMAMAKTIDSTFONO-----TRL 770
Db 1024 GLASRLVEMGRAGAFGLPVYVEIRPHSGMKGKRELIVLASNGETKEHQPCKMEDL 1083
QY 771 VTIIDGLACEDQKYLQMLDTRY--LFSKGFIAIFASDPIIITKAINQNLNSVLDSN 828
Db 1084 AELLGMD--BELDSPEELGTRICRITK-DFQYFA-----VYSRIKQESNOIGPEG 1135
QY 829 INGHYMRNIYLPFLNSRGLSNARKFLY--TSAITNGITGSDTTGTQ-----EDTDRV 882
Db 1136 I-----LSTTYPLVQASPEGLTRIRIVGLOAQVPEETVKI 1175
QY 883 SONSIGEMTKLSKTALNRDITY--RRROMORTITROMSFDLTKL--VTEDMFSDIP 937
Db 1176 -----LGNKATFSPITYVEPRRKHFKPIT--MTIYVPPSGGVSNGKGDATP 1223
QY 938 QTRMLNIIVSYTG-----RLIRANOITFNMDRLASWINTLEQMYRTSWL-----I 984
Db 1224 NL--RL--CSITGTSPAWMEDITGTPPLFIKDCVSFTNVASARF-----MLADCHOV 1274
QY 985 LYLEETEGLPOMTLTKM-YERISKNIPTTKDVEPLEIDODINF----- 1029
Db 1275 L--ETVGLASQILRELICVYMAKVFYFANTNP--VESLSCFCMTDRDVKLEOQ 1328
QY 1030 -----EVLSSRTPVLVADVKTPLPCYVNLDPKLR--ELIADYRAARE-----QINIG 1077
Db 1329 ENFEVARSKDIEVLEKPI--YVDCYGNLAPLFRKGGOQLVFNFSKRENLRPSIKIRD 1386
QY 1078 LAYPLP--LHSGPPRPPSGYQSPASVCSASFNGPFGGVVSPQPHSSYVSGLSGPQ 1133
Db 1387 TSOECCGRLSFLKE--PKTTKGLPQTA-VCN-----LNTILPA 1421
QY 1134 HPFNRAAVPATGSSLL-----SSMTVDVYCK--LROTEGLDQMMPOYCTTIK 1182
Db 1422 HKKAKKADROSFASLARKRYSTLTPBSMSPQSPCERDTRMAIVAD-HLGSLWELAR 1480
QY 1183 KANINGRYVLSQCNIDELKEMANF--GDW-----HLFSWVLEMRSESQVVP 1229
Db 1481 ELNFSYDEINOIRVENPNSLISQSFMLKKVNTDGNKATTDALTSVLTIINRDIYTL 1340
QY 1230 EDP-----RFLNENS--SAPVPHGESARRSHHTLPLTELSSQPPYTLNFSFEE 1276
Db 1541 EGPFLDYGNISGTRSFADENNVPDPV-----DGH--PSFOVELTEPMGLMYT--P 1587
QY 1277 LNTIGLDEGAPRHSNLSMOSQTRTPSLSS-----LNSDSSSTELS 1317
Db 1588 PNPPOQOD--HSDISISPFRTPSRLDGLVPSOGNIHPTGPPVYTAEDTSLSDS 1644
QY 1318 KLVDKV--QAEYRDAYREYIAQMSQLEGTSSTIG-----RSSP--HSTYVIGS 1365
Db 1645 KMDOSVYVTPADPLDVDESQLDKDQSECAQCMASVPGIIPNDGRQAEPLRPQTRKVGMS 1704

```

```

QY 1366 SSGSIHSTLEQERKKEGELKQEDGRKSFIMKRGVYIDYSSSGVSTNEASPLDPITEDE 1425
Db 1705 SE-----QOERKSG-----PDEVEDKY 1724
QY 1426 KSDSGSKLPGKSSSRPSLFDTLKKGGLRYOKLPSDEDESGTGRVQITPHCSKI 1485
Db 1725 K-----SLFE-DIQE--EVEAEEMEDQOAMINRVO-----RAEL 1758
QY 1486 RTRKLRKARECASPOESAEPLRTFLIKAEYSDLLDKKSSDSGSVSN----- 1536
Db 1759 AMSLSAQWQETPSGSLFSAQAR-----LTGGLDLRLDSDSQANDSTSYLTGE 1811
QY 1537 ---ESSPNHSLHNEADDQLEKANTIELEDESGSKRGK-----PSLSGLDPIIA 1586
Db 1812 GKTEANGNHT-----AEVIEAKAKYPFEPSONDICKOSIKENLKRKTHGCGTEEPV-- 1864
QY 1587 RMSICSEDKKSPSECS--LIASSPESWPACQAYVNLNPTSTVTLNNTPAPNANOF 1644
Db 1865 --SPLTAYQKSLBETSKLYEDAPKPCVPYMKKM-----TRTADGKARLNL 1910
QY 1645 DEIEG 1649
Db 1911 QEEBG 1915

```

```

RESULT 13
ID 013768
AC 013768; PRELIMINARY; PRT; 1719 AA.
DT 01-NOV-1996 (TEMBLrel, 01, Created)
DY 01-NOV-1996 (TEMBLrel, 01, last sequence update)
DE 01-JUN-2002 (TEMBLrel, 21, last annotation update)
DT Alt. ankyrin (variant 2.2).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=HEMATOPOIETIC;
RX MEDLINE=90158830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RT control proteins."
RL Nature 344:36-42(1990).
DR EMBL: X16609; CAA34611.1;
DR HSSP: Q00420; IAKC.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; ZUS.
DR Pfam: PF000023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ank; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 22.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZUS; 1.
DR PROSITE: PS50088; ANK_REPEAT; 20.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1719 AA; 189011 MW; F63465D16D975CBF CRC64;

Query Match 6.2%; Score 545; DB 4; Length 1719;
Best Local Similarity 20.6%; Pfam No. 4,4e-26;
Matches 376; Conservative 270; Mismatches 604; Indels 574; Gaps 71;
QY 21 LKALEKCDVDNERECGOTPLMLAEOGNEVIELKNGANCLDLDMMFTLISAK 80
Db 94 VELVNYGANNVAQSQKGFPLYMAQENHLEVAKFLLENGANONVATDESFPLVALQ 153

```

```

QY 81 EGH----- 83
DB 154 QGHENVAHLINJCTGKVRPLAHLIARNDTRTAVALLDNDPNDVLSTGTFPLHIA 213
QY 84 -----IHYELLKSGASLEHMDGWTALMAMACYGRPDVVELLSHGANDSVGLQYS 138
DB 214 AHEENINAVOLLNBRASVETFPQNGITPLHISRRGNVIMVLLDREGAQLE-TKTDE 272
QY 139 VYRILMAGRGHADIHILLONGAKVNCSDKVTGTFVLVMAARGHLECKHILMAGADV 198
DB 273 LTPHCAARGHVRISIELLDHGAP1QAKTKNGLSPHMAAQGDHLDVYRLLQYDAED 332
QY 199 -----OEGA-----NSWTLALIVAKGTYGYSKEILK 225
DB 333 DITLDLPLVHAHGHVAVKVLDDKAKPRSRALNGTPTPLHIACKNNHVRVMDLLK 392
QY 226 RNRNVNLTGKGTALMISKEGHIEIVODLDAGTYVNIIPDRSGTFLVIGARGHVEI 285
DB 393 TGA SIDAVTESGLTPHVASFMGHLPIVKMLORGASPNVSNKVEPTPLHMAARAGHTEV 452
QY 286 VRALLQKYADIDIRGDNKTALYAVEKGNATVVRDILQCNPTDELCTGDGTEPLKATK 345
DB 453 AKTLQOKAKVAKKADDDOTPLHCAARIGHTNWKILLNNANPNLATAGHTPLHIAK 512
QY 346 MRNIEVELLDKGAQVSAVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPKNAG 405
DB 513 EGHVEVYALLEKEASQACMTKKGFTPLHMAKGVRAVELLE-RDA--HNNAG 566
QY 406 E--TPYNDCSHQ-----KSLI-----TOIFGAR-- 427
DB 567 KNGITPLHVAHNNNDIYKLLPRGSGPHSPAMNGTPTPLHIAKONOVEARSLQYGS 626
QY 428 -----HSPTEGDMIGDYSSALADILSEPTMQP-----PIC-- 462
DB 627 SANAESVQVTPPLHQAQBGAEMVALLSKQANGNKGSGTPTPLHVAQBEGHVAVDY 686
QY 463 -----VGLYAWQMSGK--SEILKLLED-EMKTEAGQOTEPFLQFQSWL 501
DB 687 LIKHGVAVDATTRMGYTPPLHVASHYGNIKLVKFLQADVNAKTKLGY--SPHLQAAQ 744
QY 502 --IVFLLLLCGGLGIVFAFPVD-----TNLAIALSLFLALYIFFIYITYGREG 553
DB 745 GHNDIVYLLKNG-----ASPNEVSSDGTPLAIARLGISVTDLVKVV----- 789
QY 554 SMMAMALSTRLARHIGYELLKFMVNPPELPEQTKALPVRFETDNRLS-SVGE 612
DB 790 TDSETSFLVSDKIR-KSPRETVDEILDVSEDEGELIS-----PKAERRSRDYDE 840
QY 613 TSLAEMIATLSDACEREFGFLATRLFR-----VFTEE-SQKKKKWKTCCLPSFY 662
DB 841 KELLDYFVKLDQVE-----SPAIPRIPCAMPEVIRSEBQEQASKVEDDSLIPS-- 892
QY 663 IFLEFYOCITAGITLAIIFRVDPRKHLTVNALISASVGLAVLNCRTMWOVLDSLNS 722
DB 893 -----SPATEISDNI-SPVASPVHTGLVAF--AVDARGGS 925
QY 723 QRRKLSAASKLKLKSEGFMKVLKCEVLMARMAKTIIDFSTNOFRVYIIGDLACBQ 782
DB 926 MRGSRHN-----GLRVVIP-----PRCAAPTITTCRLV----- 954
QY 783 DKVLOMLDTPVRLVFSKGFPAIFASDPHIIKAINONLNSVLRDSNNGHDMYRN-IVHL 841
DB 955 --FKOKLSTPPPL-----AEEBGLASRIIALGPTGAQFSPYIVEI 993
QY 842 PVFL-NSRGLSNARKFLVYTSATNGDITCSPTTGTQEDTDKRVQNSIGETKIGSTALN 900
DB 994 PHRASHRG--DRELVLVASEGVS-----WEHRSRYGESYLDQITN-GMDEELG 1041
QY 901 RRDYRRORQRTTROMSFDLTKLIVT--EDWESDISQTRRLNITVSVTGLLRAN 957
DB 1042 SLELEKKRVCRIIT--TDFPLKVFVMSRCLQD-YDITGEG-----GSLSKVLPLY 1091
QY 958 QITFNMDRLASWINLLEQWMPYRTSWILLYLEETEGLPDQMTLK-----TWYER 1005

```

```

DB 1092 QATPEPNNVTKRVKALQ-----AQVPDELVTKLLGNQATFSPVTVYEP 1137
QY 1006 ISK-----NITPTDVEBLEIDODINREVF-----LSSRTPLVLA 1042
DB 1138 RRRFRRP1GLR1PLPPTDNRDSGEDTSLRLCSVIGTQDQAWMEDTGTCKLYA 1197
QY 1043 RDVKTFLPCTVLDKLEITIDVRAAREQINIGGLAVPPLHLEGPRRPSGVSOPASV 1102
DB 1198 MECANF-----TNV--SARFWSLDCRPTAEVNAFLLYKEL-----TAPYNAKF 1242
QY 1103 CSSASENGEPFGVVSPOPHSSYSYGLSGPOHPFYNRAAVPATGSSLLSSMTVYVCEK 1162
DB 1243 VIFAKMDPRBQ-----MAMNPGD-WHLFRSVMLEMRVSQVPEDEPR----- 1269
QY 1163 LKQ-----LEG-----LDONMPOYCTTIKAN-----INGVL 1191
DB 1270 LEQHNFEVARSRDIEVLEGMSLPAELSGNLYP-----VKRAAORSFHFOSPRENLA 1324
QY 1192 SQCNIDELKE-----MAMNPGD-WHLFRSVMLEMRVSQVPEDEPR----- 1233
DB 1325 MGVKVRDSSREPGSGSLPLKRAMKEDYOHILCHLNTMPCAKGSGAEDRRTPPLAL 1384
QY 1234 ---FLNENSSAPVPHGESAR-----SHTLPLTELSSQTPYTLNFSFELNTGLD-- 1283
DB 1385 RYSLISESTPGSLSTGEQEMKAVISEHLGLSMALARE--LOFSVEDINIRVENP 1440
QY 1284 ----EGAPRHSNLSMOSQTRTPPLSLNSODSSIEIKLNDKVA-----EVRD 1329
DB 1441 NSLIDQSVALLNL-WVIRGQANMENLYTALQSDREIYVMLEGGSGRQNRNLPDRR 1499
QY 1330 AVREYIAQMSOLEG-----TGSSTISGRSPSTHYIGQSSGSGSHSTLEQERKE 1382
DB 1500 TBRDYSLSPOQNGHQRQARITHSPYVS-QYTERSGORLQDMDWDGSIYGLD--AAQ 1556
QY 1383 GELKQE--DGRKSTL-----MKRGVIDYSSGVSTNEASPLDPTDEKSDQSGS 1432
DB 1557 GSMOEYVTCGPHSPFGSTMTPEGLRPGSGOEYKVLVSEBHTWTEQPEASSQADRRR 1616
QY 1433 KLLPGKSSERSLPOTDILKLGGLRYOKLPSD-----DESGTRVQITPHCSMI 1485
DB 1617 QCGQEOVOEAKNTTQVVO-----GNEFONITGEQVTEBQTFDEGN-----I 1660
QY 1486 RTKRLKAK--ORECAS--POEH 1503
DB 1661 VTKKIIRKVVRIQIDLSMDAAQEH 1684

```

RESULT 14

070511 PRELIMINARY; PRT; 2622 AA.

AC 070511:

DT 01-ANG-1998 (TREMblrel, 07, Created)

DT 01-AUG-1999 (TREMblrel, 11, last sequence update)

DT 01-DEC-2001 (TREMblrel, 19, last annotation update)

DE 270 kDa ANKYRIN G Isoform (ANKYRING) (Fragment).

GN ANK3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=98417633; PubMed=9744885;

RA Zhang X., Bennett V.

RT "Restriction of 460/270-kD ankyrin G to axon proximal segments requires multiple ankyrin G-specific domains."

RL J. Cell Biol. 142:1571-1581(1998).

[2]

RP SEQUENCE FROM N.A.

RA Carpenter S.S., Zhang X.

RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.

[3]





Db 933 DSMMEELVPSKEQHLLFTREFDSDSLRHYSMAADTLIDNNVLYSSPHHSFLVSPWDA 992  
QY 710 RTWMOVLDSLNSORRRLHSAASKLHLKSEGFMKVL---KCEVELMARNAKTIDSFQ 765  
Db 993 RG-----GSMRGSRRHG-----MILIPRKCJAP-----TR 1019  
QY 766 NOTRLVYIIDGDACEOKVLOMLDTVAVLFSKGFIAIFPSDHIITKAINQNLNSTR 825  
Db 1020 ITCRLV-----KRRKLAN-----PPPH---GRRGISTRLV 1047  
QY 826 DSNINGHYMRN-IYHLPVFLNSRGLSNARKFLVTSATNGDITCSDFGTQEDTDRRYSQ 884  
Db 1048 EMGPAGAQFLGVYIYIHPFGSMRG--KERELIYLRSENGE-----TWKEHQFD---SK 1096  
QY 885 NSLGEMTKL--GSKTALNRDQTYRRRQMQRTITROMSFDLTKLLVTEDEWESDISPQTMRR 942  
Db 1097 NE--DLTEILNGMEBELDSPELGKRRICRIITKDF-----POYFAVVS--RIKQ 1142  
QY 943 LLNIYSVTGRLLRAN-----QITFMWRLASMINLTEQWPRYRTSMILLYLEETGLPDQM 997  
Db 1143 ESNQIGPBGILISITVPLVQASFPEGALTKRIRVGLD-----AQPVPDEI 1188  
QY 998 TLKTM-----YERISKNI-----TTKVEPLEI 1022  
Db 1189 VKKILGNKATFSPITVPEPRRRKFKPIITMTIPVPPSGEGVSNYGKGDTPNLRLLCSI 1248  
QY 1023 DG-----DIRNEVLSKRTPLVARDVKTLPCTVNDPKIRELIADYRAARQIN 1074  
Db 1249 TGGTSPAQWEDI-----TGTPLTFIKDCVSF---TTNV--SAREWLADCHQVLETVG 1296  
QY 1075 ICGLAYPPLPHEGPPRPGSPGVSQASVCSASFNGPPGVSPQPHSSYSGLSGPOH 1134  
Db 1297 LATQIYREL-----ICV-----PYMAKFV-----VFAKMDP-- 1323  
QY 1135 PTYNRAAVPATGSSLLSSMTVDVYCEKLRQIEGLDQNMPOYCTTIKANINGRVLSQC 1194  
Db 1324 -----VSSLRCFCMTDDKVDKLTLEQOE----- 1346  
QY 1195 NIDELAKKEMANFGMHLFRSMVLEMRSEVQVPEDEPRFLN--ENSSAPVPHGE-----S 1248  
Db 1347 NEEV-----ARSKDIEVLEKPIYVDYCGNLAPLTKGGQOVLFN 1386  
QY 1249 ARRSHTELP-----LTELSSQTPYTLNFSFELNLTGLDEGAPRHSNLSMOSQTRTPSL 1304  
Db 1387 FYSEKENRLPFIRKINDISOBERCGSLFLKERKTKGLPOTAVCNLNTLPAHKETES- 1445  
QY 1305 SLSNQSSEIETSKLTDKVAEYRDARE---YIAQMSQLEGGTSSSTISGRSSPHSTYY 1361  
Db 1446 ---DODEIE---KTDROSFASIALRKRYSYLTPEGMIERSTGAT---RSLP-TTY- 1492  
QY 1362 IGSSSSGSIHSTLEQERKGEKQEDGRKSFMLKRGDVIDYSSSGVSTNEASPLDPIT 1421  
Db 1493 -----SYKPFSTRYQSWTAPITVGPAPKS-----GFTSLSSSSNTPSASPLKSIW 1541  
QY 1422 EDEKSDSGSKLPGKKSSEPSLFOJDLKLGGLRYQKLPDEDESGTGRVQITPHC 1481  
Db 1542 -----SVSTPSPIKSTL-----GASTTSSVKISDV 1567  
QY 1482 SKMITKRLAKQREKASPOEHSAPITFIKNEY---LSDALDK-----KDS 1528  
Db 1568 ASPIRSLRT-----MSSPIKTIVVQSOPRYNIQVSSGTLARAPAVTEATPLKGL 1614  
QY 1529 SDGVRSMESSPNHSLHNEADDQLEKANLIELEDEHSGKRGMPHS----- 1576  
Db 1615 ASNSTFSSRTSP-----VTTAGSLERSSTITMTPPASPKSNINNYSSSLPFKSIITSAA 1668  
QY 1577 -----LSGLDPIIARMSICSEDEKSPSECSLASSPEESMPA----- 1614  
Db 1669 PLISSPLKSVSPVAKRDVYISSAKITMASS---LSSPVKOMPGHAEVALVNGSISPLKY 1725  
QY 1615 COKAYNLNRTBSTVTLNRRN-TAPTNRANO---NEDEIGIRETSOVT---LRPGSPN 1665

Db 1726 ASSSTLINGKATATLQEKISSATNSVSVSAATDVEKVFSTTAMPFSPLRSVSNA 1785

QY 1666 PTAQVQENLKSMAHRSQSRSSITRLSKDASELHAASSEST 1705

Db 1786 PSAFQSLRTPSAS-----ALYTSL---GSSISATTSSVT 1816

Search completed: July 1, 2003, 14:40:08  
Job time : 109 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:04 ; Search time 16 Seconds  
(without alignments)  
4445.744 Million cell updates/sec

Title: US-10-021-571-4  
Perfect score: 8884  
Sequence: 1 MSYLISQSYINVEENIPA.....LHAASSESTGFEERESTL 1715

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555.5	6.3	1880	ANK1_HUMAN	P16157 homo sapien
2	539	6.1	3924	ANK2_HUMAN	O01484 homo sapien
3	525	5.9	1862	ANK1_MOUSE	O02357 mus musculu
4	521	5.9	4377	ANK3_HUMAN	O12955 homo sapien
5	458	5.2	768	YB23_HUMAN	O9u117 homo sapien
6	438.5	4.9	1059	Y379_HUMAN	O15084 homo sapien
7	409.5	4.6	1431	DAPK_HUMAN	P53355 homo sapien
8	376.5	4.2	1327	TNKL_HUMAN	O05271 homo sapien
9	375.5	4.2	323	ANKH_CHRVI	O06527 chromatiu
10	373.5	4.2	587	ASB2_HUMAN	O96q27 homo sapien
11	370	4.2	1166	TNKL2_HUMAN	O9h2k2 homo sapien
12	353	4.0	741	RNS5A_HUMAN	O05823 homo sapien
13	351	4.0	692	ANK6_HUMAN	O9y284 homo sapien
14	348.5	3.9	832	ANK3_MOUSE	P57078 homo sapien
15	348	3.9	735	RNS4_MOUSE	O05921 mus musculu
16	330.5	3.7	518	ASB3_HUMAN	O9y575 homo sapien
17	328	3.7	1401	LATA_LATMA	P23631 latroductu
18	324.5	3.6	525	ASB3_MOUSE	O9w772 mus musculu
19	318.5	3.6	747	V222_FOPPV	O9j513 fowlpox vir
20	318	3.6	583	AS15_MOUSE	O8v8h6 mus musculu
21	309	3.5	596	V244_FOPPV	O9j546 fowlpox vir
22	307	3.5	668	V244_FOPPV	O9j546 fowlpox vir
23	304.5	3.4	542	V155_FOPPV	O9j546 fowlpox vir
24	302	3.4	436	V245_FOPPV	O9j445 fowlpox vir
25	284	3.2	1083	V1L2_YEAST	P40460 saccharomyc
26	278	3.1	525	V228_FOPPV	O9j507 fowlpox vir
27	275	3.1	592	V246_FOPPV	O9j444 fowlpox vir
28	267	3.0	603	V162_FOPPV	O9j569 fowlpox vir
29	264.5	3.0	656	FEM1_CAEL	P17221 caenorhabdi
30	259	2.9	433	AS14_MOUSE	O8v8h7 mus musculu
31	254.5	2.9	434	AS15_HUMAN	O8wxk1 homo sapien
32	252	2.8	642	YAZA_SCHPO	O09701 schizosacch
33	242.5	2.7	429	AS10_HUMAN	O8wx13 homo sapien

34	240.5	2.7	278	1	AS13_HUMAN	O8wxk3 homo sapien
35	240.5	2.7	776	1	ANK5_HUMAN	O9nu02 homo sapien
36	239	2.7	434	1	V023_FOPPV	O9j5h8 fowlpox vir
37	235	2.6	333	1	ANK2_HUMAN	O9g2v1 homo sapien
38	233.5	2.6	2703	1	NOTC_DROME	P07207 drosophila
39	232.5	2.6	578	1	V022_FOPPV	O9j5h9 fowlpox vir
40	231.5	2.6	328	1	ANK2_MOUSE	O9wv06 mus musculu
41	231.5	2.6	437	1	V014_FOPPV	O9j517 fowlpox vir
42	231.5	2.6	461	1	V218_FOPPV	O9j517 fowlpox vir
43	230	2.6	226	1	PSDA_HUMAN	O75832 homo sapien
44	229.5	2.6	1066	1	NUC2_NEUCR	O01317 neurospora
45	229	2.6	231	1	PSDA_RAT	O922x3 rattus norv

## ALIGNMENTS

RESULT 1  
ID ANK1\_HUMAN STANDARD: PRT: 1880 AA.

AC P16157;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).  
GN ANK1 OR ANK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.  
RC TISSUE=Hematopoietic;  
RX MEDLINE=90158350; PubMed=2137557;  
RA Lux S.E., John K.M., Bennett V.;  
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins.";  
RT Nature 344:36-42(1990).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=90175370; PubMed=1689849;  
RX Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,  
RA Cheung M.C., Kan Y.W., Palek J.;  
RT "cDNA sequence for human erythrocyte ankyrin.";  
RT Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).  
RN [3]  
RP VARIANT HS ILE-462.  
RX MEDLINE=96225450; PubMed=8640229;  
RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,  
RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,  
RA Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;  
RT "Ankyrin-1 mutations are a major cause of dominant and recessive hereditary spherocytosis.";  
RT Nat. Genet. 13:214-218(1996).  
CC -1- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN G85, AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.  
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN; THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC PLASMA MEMBRANE.  
CC -1- ALTERNATIVE PRODUCTS: At least 3 isoforms; 1/2.1 (shown here), 2/2.2 and 3, are produced by alternative splicing.  
CC -1- PTM: REGULATED BY PHOSPHORYLATION.  
CC -1- PTM: ACYLATED BY PALMITIC ACID GROUP(S).  
CC -1- DISEASE: Defects in ANK1 are the cause of dominant and recessive hereditary spherocytosis (HS).  
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.  
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.  
CC



	DR	EMBL; A18609; CAA34510.1;	-.
	DR	EMBL; M28880; AAA51732.1;	-.
	DR	PIR; S08275; SJHUK.	
	DR	PIR; A35049; A35049.	
	DR	HSSP; Q00420; JAMC.	
	DR	GeneW; HGNC:492; ANK1.	
	DR	MIM; 182900;	--
	DR	InterPro; IPR002110; ANK.	
	DR	InterPro; IPR000488; Death.	
	DR	InterPro; IPR000906; ZU5.	
	DR	pfam; PF00023; ank_24.	
	DR	pfam; PF00531; death_1.	
	DR	pfam; PF00791; ZU5; 1.	
	DR	PRINTS; PR01415; ANKYRN.	
	DR	SMART; SM00248; ANK_22.	
	DR	SMART; SM00005; DEATH; 1.	
	DR	SMART; SM00218; ZU5; 1.	
	DR	PROSITE; PS50086; ANK_REPEAT; 20.	
	DR	PROSITE; PS50297; ANK_REP_REGION; 1.	
	DR	Cyoscalecom; Alternative splicing; Repeat; ANK repeat;	
	KV	Phosphorylation; Lipoprotein; Disease mutation; Elliptocytosis;	
	KW	Polymorphism.	
FT	FT	INIT_MET	0
FT	FT	DOMAIN	1
FT	FT	DOMAIN	827
FT	FT	DOMAIN	1381
FT	FT	DOMAIN	1382
FT	FT	DOMAIN	1880
FT	FT	REPEAT	43
FT	FT	REPEAT	76
FT	FT	REPEAT	109
FT	FT	REPEAT	142
FT	FT	REPEAT	173
FT	FT	REPEAT	204
FT	FT	REPEAT	237
FT	FT	REPEAT	270
FT	FT	REPEAT	303
FT	FT	REPEAT	336
FT	FT	REPEAT	369
FT	FT	REPEAT	402
FT	FT	REPEAT	435
FT	FT	REPEAT	468
FT	FT	REPEAT	501
FT	FT	REPEAT	534
FT	FT	REPEAT	567
FT	FT	REPEAT	600
FT	FT	REPEAT	633
FT	FT	REPEAT	666
FT	FT	REPEAT	699
FT	FT	REPEAT	732
FT	FT	REPEAT	765
FT	FT	DOMAIN	1402
FT	FT	VARSPLIC	1512
FT	FT	VARSPLIC	1874
FT	FT	VARSPLIC	1849
VT	VT	VARIANT	20
VT	VT	VARIANT	462
VT	VT	VARIANT	618
VT	VT	VARIANT	618

89 KDA DOMAIN (ANION EXCHANGE PROTEIN BINDING DOMAIN).  
62 KDA DOMAIN (SPECTRIN BINDING DOMAIN).  
55 KDA REGULATORY DOMAIN (REGULATES THE BINDING OF AKRYNIN TO SPECTRIN AND THE BAND 3 PROTEIN).  
ANK 1.  
ANK 2.  
ANK 3.  
ANK 4.  
ANK 5.  
ANK 6.  
ANK 7.  
ANK 8.  
ANK 9.  
ANK 10.  
ANK 11.  
ANK 12.  
ANK 13.  
ANK 14.  
ANK 15.  
ANK 16.  
ANK 17.  
ANK 18.  
ANK 19.  
ANK 20.  
ANK 21.  
ANK 22.  
ANK 23.  
DEATH.  
MISSING (IN ISOFORM 2).  
H -> D (IN ISOFORM 2).  
TVGQDPELSELEVDIDYPAKHSHDNSTPNP -> ELGRS  
GLQDLIEGRKAQIYARASLRKKOQ (IN ISOFORM 3).  
R -> T.  
. /FtId=VAR\_000595.  
V -> I (IN HS).  
. /FtId=VAR\_000596.  
R -> H (IN BRUGGEN).

[illegible]



DR HSP: P42771; 1DC2.  
 DR Genew; HGNC:493; ANK2.  
 DR MIM: 106410; -  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR000906; ZUS.  
 DR Pfam; PF00023; ank; 24.  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00791; ZUS; 1.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 21.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00218; ZUS; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 20.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 DR Cytoskeleton; Alternative splicing; Repeat; ANK repeat;  
 KW Phosphorylation.  
 FT REPEAT 63 92 ANK 1.  
 FT REPEAT 96 125 ANK 2.  
 FT REPEAT 129 158 ANK 3.  
 FT REPEAT 162 191 ANK 4.  
 FT REPEAT 193 220 ANK 5.  
 FT REPEAT 232 261 ANK 6.  
 FT REPEAT 265 294 ANK 7.  
 FT REPEAT 298 327 ANK 8.  
 FT REPEAT 331 360 ANK 9.  
 FT REPEAT 364 393 ANK 10.  
 FT REPEAT 397 426 ANK 11.  
 FT REPEAT 430 459 ANK 12.  
 FT REPEAT 463 492 ANK 13.  
 FT REPEAT 496 525 ANK 14.  
 FT REPEAT 529 558 ANK 15.  
 FT REPEAT 562 591 ANK 16.  
 FT REPEAT 595 624 ANK 17.  
 FT REPEAT 628 657 ANK 18.  
 FT REPEAT 661 690 ANK 19.  
 FT REPEAT 694 723 ANK 20.  
 FT REPEAT 727 756 ANK 21.  
 FT REPEAT 760 789 ANK 22.  
 FT REPEAT 793 822 ANK 23.  
 FT DOMAIN 1773 1950 REPEAT-RICH REGION.  
 FT REPEAT 1773 1784 REPEAT A.  
 FT REPEAT 1785 1796 REPEAT A.  
 FT REPEAT 1797 1808 REPEAT A.  
 FT REPEAT 1809 1820 REPEAT A.  
 FT REPEAT 1821 1832 REPEAT A.  
 FT REPEAT 1833 1844 REPEAT A.  
 FT REPEAT 1845 1856 REPEAT A.  
 FT REPEAT 1857 1867 REPEAT A (APPROXIMATE).  
 FT REPEAT 1868 1879 REPEAT A.  
 FT REPEAT 1880 1891 REPEAT A.  
 FT REPEAT 1892 1902 REPEAT A (APPROXIMATE).  
 FT REPEAT 1903 1914 REPEAT A.  
 FT REPEAT 1915 1926 REPEAT A.  
 FT REPEAT 1927 1938 REPEAT A.  
 FT REPEAT 1939 1950 REPEAT A.  
 FT DOMAIN 3536 3620 DEATH.  
 FT VANSPLIC 1039 1039 O -> OFGLKLTPTAPPLNEGESLVRILQGLPPTK  
 (IN ISOFORM 2).  
 VANSPLIC 1444 3528 MISSING (IN ISOFORM 2 AND ISOFORM 3).  
 FT CONFLICT 475 476 GO -> PE (IN REF. 4).  
 FT CONFLICT 971 971 I -> S (IN REF. 1).  
 FT CONFLICT 3581 3582 QY -> HA (IN REF. 1).  
 FT CONFLICT 3586 3586 I -> Y (IN REF. 1).  
 SQ SEQUENCE 3924 AA; 430337 MW; 52AC496C428E29D2 CRC64;  
 Query Match 6.18; Score 539; DB 1; Length 3924;  
 Best Local Similarity 20.18; Pred. No. 3.5e-20;  
 Matches 421; Conservative 294; Mismatches 717; Indels 662; Gaps 81;

Db 73 KEGHVGVLQVDELLRGSSVDSATKKGNFALHIASLAGAEVYVLYKEGANINAAQSONGPT 132  
 QY 74 ALISAKKEGVHIVIEELLKCGVN-----LEHRDNG--- 103  
 Db 133 PLYMAAOENHIDVKKYLENGANOSTATEDGETPLAVALQGHNAVALLENDKRGKVR 192  
 QY 104 -----GWTALMMACTKRTDYVELL 123  
 Db 193 LPALHTAARKDDTKSALLLQNDHNAVDOSKKMNVNTTESGFTPLHTAHYGNVNAVATLL 252  
 QY 124 LSHGANPSYGLQSYVPIIIMAGRGHADIHLLQNGAKVCSDKYGTPLMAARKG- 182  
 Db 253 LNRGAADVFTA-RNGITPLHVASKRGNTMMVLLDRGQIDAKTRDGLTPLHCARSH 311  
 QY 183 -----HLECVKLLMGADVDOEGANSMTALIV 210  
 Db 312 DQVVELLEBGAFLARTKNGSLPLMAAQGDHVECKHLLQKAPVDVDTLYLTLAHV 371  
 QY 211 AVKGYTOSVKEITKNPNVNLTDKDNALMASKE----- 247  
 Db 372 AAHCGHYRTKLLDKRANPNARALNGFTPLHTACKKRIKYMELLYKGASIOAITSEG 431  
 QY 248 -----GATEIVQDLLDAGTYVNIPIRSGDFVLGAVRGHVEIYRALLQRYADID 297  
 Db 432 LRFPIHVAAFMGHILNIVLLLLQNGASPDVYINIGETALHMAARAGQEVYRCLLRGALVD 491  
 QY 298 IROGDKKALYMAVEGNATWVRDIIQCPDIEICKDEETPLIKATKRNIEVELLD 357  
 Db 492 ARAREQTPHLIASRLGKTEIVQLQNHANHPAATNGTYPHLHSARGQVDAVSYLE 551  
 QY 358 KGAQYAVNDKGDPTPLHIAIRGRSRKLAELLR-----NPRD 394  
 Db 552 AGAAHLATKKGFTPLHVAAKYGSIDYATKLLQRRAADSAKNGTPLHVAHNDQKV 611  
 QY 395 GRLL-----YRPNKAGETPINIDC-SHQKSLIOI-FGAR-----HLSPT 432  
 Db 612 ALLLEKGAHPATANGYTPHLIAKKRQMOIASLTLNLYGAEFTNIVTKOGVTPPLHASQ 671  
 QY 433 ETDGDML-----GYDLYSSALADILSEPTMQPICOVGLYAOGSGKSLKLLKLEDEM- 484  
 Db 672 EGTIDVNTLLDKGANIHMTSGSLTS-----LHLAAQ-----EDKXNV 710  
 QY 485 -----KTFAGQOIEPLFOPSWLIVPLTLILC--GGILGL-FAFTYHPNGLAVSISFPL 536  
 Db 711 ADILTFHGADQDAHTKLGTYPLIV-----ACHYGVKVMVNLKQGANVNATKNGYPL 765  
 QY 537 -----LYIFPIYVYFGRGREGSWANAVLSRLAHGYELL-LKIM----- 579  
 Db 766 HQAAQGHHTIINVLLQHGAKPNATYANGNTALA--IAKRLGISVVDLKVVEEVTT 823  
 QY 580 -----FVNPELPEQTTKALP-VRF-F 600  
 Db 824 TITTEKHLNVPETITEVLDVSDGDDTMTGDCGEYLRPDLKELGDDSLPSQIDG 883  
 QY 601 TDYNRLSVGGET-SL-----AEMIALTDACERE 629  
 Db 884 MNYLRYSLESGRSDSLRSFSSDSRSHLSHASYLRDSAVMDSVLPSHQVSLAKEARN 943  
 QY 630 RGFIAFLRFRVKTEDTQCKKKKTKCCPLSPYITLF-----IIGCIISGTLIAIRV 683  
 Db 944 SYRLS-----WGTELDNVALSSSPIHSGFLVITPMVADARGAMGRCHNGLRIT--I 993  
 QY 684 DKPHLTNVAVLISIASVGLAFVLCRTMWQVLDLSNSORKL-----HMAASKL 734  
 Db 994 PRKRCAPF-----RYVCRL-----YKRRRLATMPMWBGELASRL 1030  
 QY 735 HKIKSEG-----FMKVLKCEVELMARNAKTIDSETON-----QTRLVIIIDG 776  
 Db 1031 IGVGSGAQFLGPIVIEIPHFALRGKREHLVLSENGDSMKHFCDYTEDLEINELNG 1090  
 QY 777 LDACEDKVLQMLDVRYVLFSGKPIALFASPPH--IILKAINQNLNSVLDNSINCHDY 834  
 Db 1091 MD-----EVLDSPEDELEKKRICRITTRDFPOYFAVVSRIKO-----DSMLIGPE 1134





Db 1730 GEPALOEYEQVLY---STREHVGREPCTSPKAGKEPSLMAFESAFS 1776

RESULT 4  
ID ANK3\_HUMAN STANDARD; PRT; 4377 AA.  
AC 012955;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ankyrin 3 (ANK-3) (Ankyrin G).  
GN ANK3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
NC TISSUE=Brain stem;  
RX MEDLINE=95138209; PubMed=7836469;  
RA Kordell E., Lambert S., Bennett V.;  
RT Ankyrin. A new ankyrin gene with neural-specific isoforms localized  
RT at the axonal initial segment and node of Ranvier.";  
RL J. Biol. Chem. 270:2352-2359(1995).  
CC -1- FUNCTION: Membrane-cytoskeleton linker.  
CC -1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by  
CC alternative splicing.  
CC -1- TISSUE SPECIFICITY: Expressed in brain and other tissues.  
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.  
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
-----  
DR EMBL: U13616; AAA64834.1; -  
DR HSSP: P5273; IB18.  
DR Genew: HGNC:494; ANK3.  
DR MIM: 600465; -  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000488; Death.  
DR InterPro: IPR000906; Z05.  
DR Pfam: PF00023; ank; 24.  
DR Pfam: PF00531; death; 1.  
DR Pfam: PF00791; Z05; 1.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00248; ANK; 21.  
DR SMART: SM00005; DEATH; 1.  
DR SMART: SM00218; Z05; 1.  
DR PROSITE: PS50088; ANK\_REPEAT; 21.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE: PS50017; DEATH\_DOMAIN; 1.  
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.  
FT REPEAT 73 102 ANK 1.  
FT REPEAT 106 135 ANK 2.  
FT REPEAT 139 168 ANK 3.  
FT REPEAT 172 201 ANK 4.  
FT REPEAT 203 230 ANK 5.  
FT REPEAT 234 263 ANK 6.  
FT REPEAT 267 296 ANK 7.  
FT REPEAT 300 329 ANK 8.  
FT REPEAT 333 362 ANK 9.  
FT REPEAT 366 395 ANK 10.  
FT REPEAT 399 428 ANK 11.  
FT REPEAT 432 461 ANK 12.  
FT REPEAT 465 494 ANK 13.  
FT REPEAT 498 527 ANK 14.  
FT REPEAT 531 560 ANK 15.

FT REPEAT 564 593 ANK 16.  
FT REPEAT 597 626 ANK 17.  
FT REPEAT 630 659 ANK 18.  
FT REPEAT 663 692 ANK 19.  
FT REPEAT 696 725 ANK 20.  
FT REPEAT 729 758 ANK 21.  
FT REPEAT 762 791 ANK 22.  
FT REPEAT 795 825 ANK 23.  
FT DOMAIN 1519 1898 SER-RICH.  
FT 4090 4174 DEATH.  
SQ SEQUENCE 4377 AA; 480399 MW; F42379E55768B684 CRC64;  
Query Match 5.9%; Score 521; DB 1; Length 4377;  
Best Local Similarity 20.6%; Pred. No. 3.86-19;  
Matches 427; Conservative 284; Mismatches 695; Indels 668; Gaps 87;  
14 EEENIPALKLEKCDVDNERNECGOPLMAEÖGNLEVKELIKNGANCNEEDDNMT 73  
83 KEHVEVYSELLQREAVNDATKKGNTALHIASIAGAEVVKVLYTNGAVNNAÖSNGFT 142  
74 ALIASKEGHVHYEELLKCGVNLHHRDMGWTALMAACYKGRDVEILL----- 124  
143 PLVMAÖENHLEVVKFLIDNGASÖSLATEDGFTPLAVALÖÖGHÖVÖSLLENDTKGVR 202  
125 -----SHGANPSVTGL----- 135  
203 LPALHIAARKDPTKAALLÖÖNDNADVESGFTPLHIAHAYGINIVATLLNRAAVD 262  
136 ---QYVPTITMAÖRGHADIVHLLÖNGAKVNCSDRYGTPPLVMAARK----- 182  
263 FTARNDITPLHVAÖKGNANVVKLLDRGAKIDAKTRDGLPLHCGARSGEÖVEMILD 322  
183 -----HLECYHLLANGADVÖDEGANSMTALIVAVKGYÖS 219  
323 RAAPILSTKKGSLPMAÖTGDHNCVÖLLÖHNVPÖDVTNDYTLAHVAHCHGYKV 382  
220 VKEILKRNPNVNLÖDKÖGTALMAÖSK----- 247  
383 AKVLLDKKANNAALNÖFTPLHACKNRKIKVEMLLKHGASIQAVTESGFTPIHVAÖ 442  
248 -GHTEIVÖDLDAGTYVNIÖPDRSGDYVLIGAVRGHVEYRALLÖKKADIDRÖÖDKTA 306  
443 MGHVIVISÖLHNHGSPÖTÖTVNÖRGÖTALHMAÖRGAÖEVVYVÖDÖGAÖVEAKKÖDÖTP 502  
307 LYMAVÖKGNATMÖVDIÖ---CNPÖT-----EÖCT 333  
503 LHSARLÖKADIVÖÖLÖÖGASPNNAÖTSGYTPLHLSARSGHEDVAAFLDÖHGASLÖTÖT 562  
334 KÖGÖPLÖLKATÖRMÖIEVÖELLÖDKGAKVSAVÖDKGÖPTPLHIAÖTRGRSÖKLAEÖLLRNPK 393  
563 KKGFTPLHVAÖKYÖGLÖEVANÖLLÖKSASPDAGSGÖLPLHVAHAYNÖKÖVALLÖL---D 619  
394 DÖRLÖYRNKAGÖTYNIDC--SHÖKSILÖTÖ--FGAR-----HSPÖETÖGDÖM 438  
620 ÖGASPHAAKNGYÖPLHIAÖKNÖMÖIATÖLÖEYÖGADANAVTÖRÖIASVHÖLAAÖEGHVDÖ 679  
439 LÖYDYSALADYIÖSEPTÖMÖP-----PÖCYÖLY 466  
680 VSLÖLGRNANVNLÖSKSGÖLPLHIAÖEDRVNAÖEVLVNCÖAHYADÖTKMÖYTPLHVGCH 739  
467 AÖWGSÖGK--SFÖLK-KLEDEMKÖTÖAGÖÖIEÖPLÖFSWÖLIVÖTLÖLÖGÖ---LÖLÖFAF 519  
740 ---YÖNÖIKÖIVNÖLLÖÖSAKAVNAÖTKNG--YÖPLHÖGAÖÖ-----ÖGHÖHIIÖVNLÖN 785  
520 TVHÖN-----LÖIÖVSLÖFALÖLYIFÖIVYÖFGÖRGÖSGÖSMNNAVÖLSTÖLANH-I 569  
786 NASPÖELTVNÖNTALÖIÖARLÖGYISVÖDTLÖKIVT-----EÖTMÖTÖYÖTEKHÖM 834  
570 GYÖELLÖKIMÖVNPÖEL-----PÖÖTKALPÖVRFÖLTDYÖNRLÖSV--GÖÖSÖLÖA----- 616  
835 NVÖETÖNEVÖLDSÖDEÖVÖKANAPÖ-----MÖSDÖEYISÖVEÖGÖDANÖMTÖDTÖKYL 884  
617 --EMÖIATÖLSDACÖERÖFGLÖATÖRL-----FÖVÖFKÖED--TÖGÖKKÖWKÖTÖCLPÖSVIÖELF 666

```

Db      885 GPDLKELGDDSLPAEGYMGFSLGARASLSRFSFSDRSTYTLNRRSYAR-----DSMIDEL 940
Qy      667 IIGCIISGTTTLAIRVPD-KHLT-----VNAVLSIASVGLAFVLCRTMVOYLD 717
Db      941 IVPSEKQHLTTFREDSDSLRHYSWAADTLDNVLNYSPIHSGFLVSPVARG-----994
Qy      718 SLINSQKRLHNAASKLHKLKSEGFMKVL-----KCEVELMARMAKTIDSPONQTRIVYI 773
Db      995 ---GSMGRSRHNG-----MRLITPRKCTAP-----TRITCRVL--1025
Qy      774 IDGLACQDQVLOMLDTVRVLESKGPRIALFASDPHIIKAINONLNSVLRDSNINGHD 833
Db      1026 -----KHKLAN-----PPPH-----GERGISRLVEMGPAGAQ 1055
Qy      834 YWRN-IYHLPVFLNRSGLSNARKFLYTSATNGDPCSTTGIOEDADRRVQNSIGETK 892
Db      1036 FLGPVIVELPHFGSRG--KERELIVLSENGE-----TWKEHQFD--SKNE--DLTE 1102
Qy      893 L--GSKTALNRDRTYRRROMORTYRQMSFDLTKLVTEDWFSIDISQPMRLNIVSVT 950
Db      1103 LLNGDEBELDSEELGKRIKITKD-----PQFAYVS--RIQESNOIQPE 1150
Qy      931 GRLRAN-----QISENDRLASWINTEQMPRTSWLILYLEETEGIPDMQTLKTYER 1005
Db      1151 GGLISTTVPLVQASFEGALTKRIRVGLQ-----AQPVDEIVKILGNK 1196
Qy      1006 ISKNPTKDYEPLEIDGIRNFEVLSRTPVLVANDVKVFLPCYVNLDPKRLIAD 1065
Db      1197 AT-----FSPITVEPRRRKFKPIITMTIP-----1222
Qy      1066 VRAAREQISIGLAVPLPLHEGPRAPSGYSOP--PSVCSSTFNGFAGVVSPOPHS 1123
Db      1223 -----PPSGEGVSNYKGDTPNRLRLCSIT-----GGRSP-----1254
Qy      1124 SYISGMTGPHRPFYKRGSGPARGPVLLN-----SLAND-----AVCEKLEQIGLDQSM 1173
Db      1235 AQMEDITQT-----PLTFIKDCVSEFTTNSANFMLDCHQVLETGLATOL 1301
Qy      1174 -----LPQCTTIKANINGRVLACNI-----DELK--EMMANFGDMHLPSTYLEM 1220
Db      1302 YRELICVYMAKFVFAKANDPVESSLRCPCMTDKDKITLQEOENTEE-----VA 1332
Qy      1221 RNAESHVVPEDPRFLSESSSGPARHGPARRA-----SHNELPHT--ELSSQTPYT 1269
Db      1353 RSKDLEVLGKRIYV-DCYGNLAPLTKGQQLVFNFFYSFKNRLPFISIKIRDTQBPGR 1411
Qy      1270 LNFSEELNTGLDGBARHNSLSMOGTRTTPSLNSLSQDSSIEISKLTDKYQAEYRD 1329
Db      1412 LSFLEKERTTKGLPQTAVCNINLTLPARKKETES-----DODDEIE--KIDRRQSPASL 1463
Qy      1330 AYRE---YIAQMSOLEGGPGSTTISGRSSPHSTY-----YMGQSSSGGSHSNLE 1376
Db      1464 ALRKRYSLTERPGMIERTGAT-----KSLP--TTSYTPFSTSTRYQSMNTAPITV-----1513
Qy      1377 QEKGDSEPKPDGRKSFMLKRGDIVYSSSVSTNDASPLDPI-----TEDEKSDSGS 1432
Db      1514 -----PGPAKSGF-----TSLSSSSNTPASPLKSIWSVSTPSPKISTGAS 1556
Qy      1433 KLDPGKSSERSLSLFQDDIKKSGLR---YQKLPSDEDESGT-----ESDNTPL-----1480
Db      1557 TTSSVKSISDVASPIRS-LRTMSSPIKTIVVSQSPYNQVSSGTLARAPAYTEAPLPLGLA 1615
Qy      1481 -----LKDKDKRAEGKVER-----VPSPE-----HSAEDIRFEI--KAKYLSDAL 1521
Db      1616 SNTFSSTSPVTYTAGSLERSSTTMTPPSPKSNITMYSSSLPFKSIITISAAPLISPL 1675
Qy      1522 -----LDKQSSDSGVRSSESSPNHSL--HNEVA-----DQSOLEKAN--LLE--1560
Db      1676 KSVSPYKSRVDYISSAKITWASSLSIPVKQMPGHABALVNGSISPLKASSSTLNGC 1735
Qy      1561 -----LDDSHSGKRGIPHLSGLDPIIARMSICSDKSPSCSLASSPEENWPAQ 1615

```

```

Db      1736 KATATLOEKISSATNSVSSVSAATDVEKFVSTTAMPFSPRLR-SYVSAP-----1786
Qy      1616 KAYLNFTPSTVTLNNSAPANRANQNFEMEGIRETSGVILRPSS-----PNPTTON 1670
Db      1787 SAFQSLTPPSAALYSLSGSSISATIS-----SYTSSITVPVSVVNVLPKALTKL 1839
Qy      1671 ENLKSMTHKRSQSS--YTRLSKDPPELHAASS 1702
Db      1840 PDSNSTKSAALLSPKITLTTEHPHPSRTS 1873

```

```

RESULT 5
YB23_HUMAN
ID YB23_HUMAN STANDARD; PRT; 768 AA.
AC Q9ULJ7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA1223 (Fragment).
GN KIAA1223
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hiroseawa M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 6:337-345(1999).
CC -1- SIMILARITY: CONTAINS AT LEAST 14 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AB033049; BAA86537.1; -
DR HSSP: PA4271; IDC2.
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ANK_14.
DR SMART: SM00248; ANK_13.
DR PROSITE: PS50088; ANK_REPEAT; 13.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT NON_TER 1
FT REPEAT 1 11
FT REPEAT 15 44 ANK 1.
FT REPEAT 48 82 ANK 2.
FT REPEAT 86 115 ANK 3.
FT REPEAT 119 148 ANK 4.
FT REPEAT 152 181 ANK 5.
FT REPEAT 185 214 ANK 6.
FT REPEAT 218 247 ANK 7.
FT REPEAT 251 280 ANK 8.
FT REPEAT 284 313 ANK 9.
FT REPEAT 317 346 ANK 10.
FT REPEAT 350 379 ANK 11.
FT REPEAT 383 412 ANK 12.
FT REPEAT 446 ANK 13.
FT SEQUENCE 768 AA; 82819 MW; 2913B69BEZDFE06D CRC64;

```

Query Match 5.2%; Score 458; DB 1; Length 768;  
 Best Local Similarity 27.9%; Pred. No. 5.8e-17;  
 Matches 122; Conservative 78; Mismatches 163; Indels 74; Gaps 4;





Db 545 TDMLSDSNKRATIS-PLHLAAVGHQALEVYOSLIDLVYRNSGRTPLDLAKFHYE 603  
 QY 504 FTLTLCGGIGLL 516  
 Db 604 CVDVLLINOASTIL 616

RESULT 7  
 DAPK\_HUMAN  
 ID DAPK\_HUMAN STANDARD; PRT; 1431 AA.  
 AC P53355;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Death-associated protein kinase 1 (EC 2.7.1.-) (DAP kinase 1).  
 OS DAPK1 OR DAPK.  
 GN Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=95129831; PubMed=7828849;  
 RA Deiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;  
 RT "Identification of a novel serine/threonine kinase and a novel 15-kD  
 RT protein as potential mediators of the gamma interferon-induced cell  
 RT death";  
 RL Genes Dev. 9:15-30(1995).  
 RP REVISIONS TO 164-171.  
 RA Feinstein E.;  
 RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.  
 CC -! FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL  
 CC DEATH.  
 CC -! PTM: AUTOPHOSPHORYLATED.  
 CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -! SIMILARITY: CONTAINS 10 ANK REPEATS.  
 CC -! SIMILARITY: CONTAINS 1 DEATH DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X76104; CA53712.1; -;  
 DR HSSP: O63450; 1A06.  
 DR GeneW; HGNC:2674; DAPK1.  
 DR MIM; 600831; -;  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000488; Death.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR Pfam; PF00023; ank; 8.  
 DR Pfam; PF00069; kinase; 1.  
 DR Pfam; PF00531; death; 1.  
 DR ProDom; PD000001; Euk\_kinase; 1.  
 DR SMART; SM00248; ANK; 7.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 6.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR Transferrase; Serine/threonine-protein kinase; Calmodulin-binding;  
 KW Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.  
 FT DOMAIN 13 266  
 FT DOMAIN 334  
 FT REPEAT 378 407 ANK 1.

FT REPEAT 411 440 ANK 2.  
 FT REPEAT 444 473 ANK 3.  
 FT REPEAT 478 507 ANK 4.  
 FT REPEAT 511 540 ANK 5.  
 FT REPEAT 544 573 ANK 6.  
 FT REPEAT 577 606 ANK 7.  
 FT REPEAT 610 639 ANK 8.  
 FT REPEAT 676 705 ANK 9.  
 FT REPEAT 716 745 ANK 10.  
 FT NP\_BIND 1313 1397 DEATH.  
 FT NP\_BIND 19 27 ATP (BY SIMILARITY).  
 FT BINDING 42 42 ATP (BY SIMILARITY).  
 FT ACT\_SITE 139 139 BY SIMILARITY.  
 FT MUTAGEN 42 42 K->A: LOSS OF ACTIVITY.  
 SQ SEQUENCE 1431 AA; 160017 MW; 9EE84811004A155B CRC64;

Query Match 4.6%; Score 409.5; DB 1; Length 1431;  
 Best Local Similarity 22.0%; Pred. No. 5,9e-14;  
 Matches 225; Conservative 149; Mismatches 338; Indels 309; Gaps 36;

QY 5 ISOSVINYVEENIPALKECK--DYDERNEGQTPMTAAEGNLEIYEIKNGA 62  
 Db 344 VMKAIHAIINDVNPGLQHLSSLNVDVNPKNHGTPLLAAGCGNLIQLIKRGS 403  
 QY 63 NCMLEDDIMWTALISAKGKHVIVEELKCGVNLHPRMGMTALMAACYKGTDYVEL 122  
 Db 404 RIDVDKGGSNAYMAARKGHDTLKFLEKNCPLDYKKSSEMALHYAARGHADVAV 463  
 QY 123 LLSHGAPSVYGLQSVYPTIIMAAARGHADIYHLLONGAKVNSDKYGTPLVMAARK 182  
 Db 464 TCASAQIDPISRTKEETPLHCAAHGYSAKALCEAGCNVNIKREGETPLTASARG 523  
 QY 183 HLECVHLLAGADVDEGANSNTALIVAKGYQSVYELLKRPNVLTPTKDGNTALM 242  
 Db 524 YHDIYVCLAEHGD-----LNACDKGHIALH 550  
 QY 243 IASKEHTEIVODLDAGTYVINIPDRSGDTVLIGARGHVEIVRALQKYDIDIRQD 302  
 Db 551 LAYRRCQMEVITLLSGCFVDYQDRHGTPLHVACKDSNMIIYALCE----- 599  
 QY 303 NKTALYAVEKGNATVRDILQCNPTETCTYDGETPLIKATKRNIEYVELLDKAKV 362  
 Db 600 -----ANCNLD--ISNRYGTRPLHANNIGLDVRYTCLMGASV 637  
 QY 363 SAVDKKDPPLHLAIGRSKLAELLNRPKGRLLYRNKAGETPPYNDCHOKSILIQ 422  
 Db 638 EALTDTGKTAEDLAREQEHVAGILARKD-----THKGLFTQQ 678  
 QY 423 IFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLVYAQMGSGKSFLLKLED 482  
 Db 679 -----LRPTQ-----NLQPRIKLKLFGHSSGSKTYTVESLK- 709  
 QY 483 EMTFRAGQIEPLPQFSMLIVLTLLCGGLGLLAFYVHPMLG----- 526  
 Db 710 -----GGLLSFPR-RRRPLSTSTNSRPPPSPLASKP 741  
 QY 527 -IAVSLFLALYIFETIVYFCGRREGESNMNAWVLSIRLAHIGYELLLKIMVNPPE 585  
 Db 742 TVSVSTNNL-----YPCENSVRSNSMPEEGILK--GMLE-----VFVAPTH 783  
 QY 586 LP-----EQTRAL-----PVRFLEFDYNRLSVSGESTLAMI 619  
 Db 784 HPHCSADQSTKAIDIONAYLNGVDFSWERSGNPVYFCYD---FAANDPISIHVV 840  
 QY 620 AILSDACEKEFG--FLATRLFRVKTEDT---QKKKKWKTCCCLPSVIFLFIITCITS 673  
 Db 841 FSLSEPEYELQNPVIFWLSFLKSLVVEEPLAFGKLR-----NPLQVVL----- 885  
 QY 674 GTTLAIFRVDKHLTVANALISVAGFLVLCRTFMQVLDLSLQSKRLHN--A 730  
 Db 886 -----VATADINVRPRGGEFGYDKD-----SLKEINRNGNDLHT 925  
 QY 731 ASKLHLKLSGEF---MKVLKCEVELMARMAKTIIDFTQNTRLVVIIDGADCEQDKVL 786

```

Db      926 SMTLFLVDAGASSGSKMKVLRNLQ--EIRQIVSVCPMTL-----CE--KIT 971
QY      787 QMLDTVAVLFSGKPFIAI--FASDPHIIKAINQNSVLRSDNSINGHDYMRNIYLVPE 844
Db      972 STLPSPWRKLNGPQLMSLOOFYVD-----VDDQLNPLASEED-----LRRIAQO--- 1015
QY      845 LNSRGISNARKFLVTSNTNDVPCSTPTGIGEPADRRVSONSIGEMTKLGSKTALNR-RD 903
Db      1016 LHSSTGEIN---IMQSEIVQDVLLLD-----PMLCTNVIGKLLSVETPRALHNRG 1063
QY      904 TYRRROMQRTTRQMFEDLTKLVTEDEWFS-DISPTMRRLNIVSVTGRLLRANOISFN 962
Db      1064 RIVVEDIQRLVPDSVDELLQILIDANDICARDLSGTM-----VDVPA-LIKTDVLHRS 1116
QY      963 W 963
Db      1117 W 1117

```

## RESULT 8

```

TNKL_HUMAN          STANDARD;          PRT; 1327 AA.
ID      TNKL_HUMAN          0952721; 095272;
AC      15-JUN-2002 (Rel. 41, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DE      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Tankyrase 1 (EC 2.4.2.30) (TNKL1) (Tankyrase I) (TNKS-1) (TRF1-
GN      interacting ankyrin-related ADP-ribose polymerase).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC      TISSUE-Testis;
RX      MEDLINE=99040105; Pubmed=9822378;
RA      Smith S., Giliat I., Schmitt A., de Lange T.;
RT      "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";
RL      Science 282:1484-1487(1998).
RN      [2]
RP      SUBCELLULAR LOCALIZATION.
RX      MEDLINE=99454782; Pubmed=10523501;
RA      Smith S., de Lange T.;
RT      "Cell cycle dependent localization of the telomeric PAPP, tankyrase,
RT      to nuclear pore complexes and centrosomes.";
RL      J. Cell Sci. 112:3649-3656(1999).
RN      [3]
RP      FUNCTION, AND PHOSPHORYLATION.
RX      MEDLINE=20556282; Pubmed=10988299;
RA      Chi N.-W., Lodish H.F.;
RT      "Tankyrase is a golgi-associated mitogen-activated protein kinase
RT      substrate that interacts with IRAP in GLUT4 vesicles.";
RL      J. Biol. Chem. 275:38437-38444(2000).
RN      [4]
RP      FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.
RX      MEDLINE=21602874; Pubmed=111739745;
RA      Cook B.D., Pyrek J.N., Chang W., Shostak G., Smith S.;
RT      "Role for the related poly(ADP-ribose) polymerases tankyrase 1 and 2
RT      at human telomeres.";
RL      Mol. Cell. Biol. 22:332-342(2002).
CC      -1- FUNCTION: May regulate vesicle trafficking and modulate the
CC      subcellular distribution of SLC2A4/GLUT4-vesicles. Has PAPP
CC      activity and can modify TRF1, and thereby contribute to the
CC      regulation of telomere length.
CC      -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose]1(N)-acceptor =
CC      nicotinamide + [ADP-D-ribose]1(N+1)-acceptor.
CC      -1- SUBUNIT: Oligomeric and associates with TNKS2. Interacts with
CC      the cytoplasmic domain of INPEP/Otase in SLC2A4/GLUT4-vesicles.
CC      Binds to the N-terminus of telomeric TRF1 via the ANK repeats.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC      with juxtanuclear SLC2A4/GLUT4-vesicles. A minor proportion is

```

```

CC      also found at nuclear pore complexes and around the pericentriolar
CC      matrix of mitotic centrosomes. During interphase, a small fraction
CC      of TNKS is found in the nucleus, associated with TRF1.
CC      -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC      produced by alternative splicing.
CC      -1- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
CC      -1- PTM: Upon insulin-stimulation, phosphorylated on serine residues
CC      by MAPK kinases.
CC      -1- PTM: ADP-ribosylated (-auto).
CC      -1- SIMILARITY: BELONGS TO THE PAPP FAMILY.
CC      -1- SIMILARITY: CONTAINS 15 ANK REPEATS.
CC      -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

```

DR      EMBL: AF082556; AAC79841.1; -
DR      EMBL: AF082557; AAC79842.1; -
DR      EMBL: AF082558; AAC79843.1; -
DR      EMBL: AF082559; AAC79844.1; -
DR      HSSP: Q00420; IAWC.
DR      Genew: HGNC:11941; TNKS.
DR      MIM: 603303; -
DR      InterPro: IPR002110; ANK.
DR      InterPro: IPR001660; SAM.
DR      Pfam: PF00023; ank; 21.
DR      Pfam: PF00536; SAM; 1.
DR      PRINTS: PR01415; ANKYRIN.
DR      SMART: SM00248; ANK; 15.
DR      SMART: SM00454; SAM; 1.
DR      PROSITE: PS50088; ANK_REPEAT; 15.
DR      PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR      PROSITE: PS50105; SAM_DOMAIN; 1.
KW      Transferrase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW      Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
KW      Phosphorylation; Alternative splicing.
FT      REPEAT 215..247
FT      REPEAT 248..280
FT      REPEAT 281..313
FT      REPEAT 368..400
FT      REPEAT 401..433
FT      REPEAT 434..466
FT      REPEAT 467..499
FT      REPEAT 500..532
FT      REPEAT 533..565
FT      REPEAT 566..598
FT      REPEAT 599..631
FT      REPEAT 632..664
FT      REPEAT 665..697
FT      REPEAT 698..730
FT      REPEAT 731..763
FT      REPEAT 764..796
FT      REPEAT 797..829
FT      REPEAT 830..862
FT      REPEAT 863..895
FT      REPEAT 896..928
FT      REPEAT 929..961
FT      REPEAT 962..994
FT      DOMAIN 1030..1089
FT      DOMAIN 1090..1176
FT      DOMAIN 1177..1327
FT      DOMAIN 1328..14
FT      DOMAIN 14..34
FT      DOMAIN 35..128
FT      DOMAIN 129..145
FT      DOMAIN 146..643
FT      VARSPLIC 644..1327
FT      VARSPLIC 1184..1184
FT      MOTAGEN 1184..1184
FT      MOTAGEN 1291..1291
FT      MOTAGEN 1291..1291
FT      MOTAGEN 1291..1291
SO      SEQUENCE 1327 AA; 142010 MW; E14DB985C710B957 CRC64;
Query Match 4.2%; Score 376.5; DB 1; Length 1327;
Best Local Similarity 22.3%; Pred. No. 3e-12;
Matches 160; Conservative 89; Mismatches 220; Indels 247; Gaps 17;

```





RN [4] SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.  
 RP TISSUE=Placenta;  
 RC MEDLINE=21443728; PubMed=11454873;  
 RA Kaminker P.G., Kilm S.-H., Taylor R.D., Zebardjian Y., Funk W.D.,  
 RT Morin G.B., Yaswen P., Campisi J.;  
 RN "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes  
 RT rapid induction of cell death upon overexpression.";  
 RL J. Biol. Chem. 276:35891-35899(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND  
 RP LNMP/OTASE.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=21661461; PubMed=11802774;  
 RA Sboadio J.I., Lodish H.F., Chi N.-W.;  
 RT "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1  
 RT (telomere-repeat-binding factor 1) and IRAP (insulin-responsive  
 RT aminopeptidase)".  
 RN Biochem. J. 361:451-459(2002).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Yin Y., Gelmann E.P.;  
 RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Tracey A.;  
 RN Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE OF 838-1151 FROM N.A.  
 RA TISSUE=Placenta;  
 RC Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,  
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,  
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
 RA Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Oshima A.;  
 RN "NEO human cDNA sequencing project".  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP FUNCTION, AND ADP-RIBOSYLATION.  
 RX MEDLINE=21602874; PubMed=11739745;  
 RA Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;  
 RT "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2  
 RT at human telomeres".  
 RL Mol. Cell. Biol. 22:332-342(2002).  
 RN [10]  
 RP FUNCTION: May regulate vesicle trafficking and modulate the  
 RP subcellular distribution of SLC2A4/Glut4-vesicles. Has PARP  
 RP activity and can modify TRF1, and thereby contribute to the  
 RP regulation of telomere length.  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose]((n)-acceptor =  
 CC nicotinamide + [ADP-D-ribose]((n+1)-acceptor.  
 CC -1- SUBUNIT: Oligomerizes and associates with TNKS. Interacts with the  
 CC cytoplasmic domain of LNMP/OTase in SLC2A4/Glut4-vesicles. Binds  
 CC to the N-terminus of Grb14 and TRF1 with its ankyrin repeat  
 CC region.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and  
 CC with juxtanuclear SLC2A4/Glut4-vesicles. Also found around the  
 CC pericentriolar matrix of mitotic centromeres. During interphase, a  
 CC small fraction of TNKS2 is found in the nucleus, associated with  
 CC TRF1.  
 CC -1- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,  
 CC liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral  
 CC blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,  
 CC ovary, small intestine, colon, mammary gland, breast and breast  
 CC carcinoma, and in common-type meningioma. Highly expressed in  
 CC fetal liver, heart and brain.  
 CC -1- PM: ADP-ribosylated (-auto).  
 CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -1- SIMILARITY: CONTAINS 15 ANK REPEATS.  
 CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

```
CC      use by non-profit institutions as long as its content is in no way  
CC      modified and this statement is not removed. Usage by and for commercial  
CC      entities requires a license agreement (see http://www.isd-sib.ch/announce/  
CC      or send an email to license@sib-sib.ch).  
-----  
DR      EMBL, AF305081; AACG5674.1; ALT_INIT.  
DR      EMBL, AF264912; AAG44694.1; -  
DR      EMBL, AF329696; AAU3463.1; -  
DR      EMBL, AF342982; AAK2581.1; -  
DR      EMBL, AF309033; AAK82330.1; -  
DR      EMBL, AF438201; AAL40795.1; -  
DR      EMBL, AL359707; CAC78760.1; -  
DR      EMBL, AK023746; BAB14665.1; ALT_INIT.  
DR      HSSP: 000420; IAWC.  
DR      Genew, HGNC:15677; TNKS2.  
DR      InterPro, IPRO02110; ANK.  
DR      InterPro, IPRO01660; SAM.  
DR      Pfam, PF00036; ank; 20.  
DR      Pfam, PF00536; SAM; 1.  
DR      PRINTS: PR01415; ANKYRIN.  
DR      SMART, SM00248; ANK; 15.  
DR      SMART, SM00454; SAM; 1.  
DR      PROSITE, PSS0088; ANK_REPEAT; 15.  
DR      PROSITE, PSS0297; ANK_REPEAT_REGION; 1.  
DR      Prosite, PSS0105; SAM_DOMAIN; 1.  
KW      Transferrase; Glycosyltransferase; NMD; Golgi stack; Telomere;  
KW      Nuclear protein; Repeat; ANK repeat; ADP-ribosylation.  
FT      REPEAT          57      89      ANK 1.  
FT      REPEAT          90      122     ANK 2.  
FT      REPEAT         123      155     ANK 3.  
FT      REPEAT         210      242     ANK 4.  
FT      REPEAT         243      275     ANK 5.  
FT      REPEAT         276      308     ANK 6.  
FT      REPEAT         363      398     ANK 7.  
FT      REPEAT         399      431     ANK 8.  
FT      REPEAT         432      464     ANK 9.  
FT      REPEAT         525      557     ANK 10.  
FT      REPEAT         558      590     ANK 11.  
FT      REPEAT         591      623     ANK 12.  
FT      REPEAT         678      710     ANK 13.  
FT      REPEAT         711      743     ANK 14.  
FT      REPEAT         744      776     ANK 15.  
FT      DOMAIN          873      936     SAM.  
FT      DOMAIN          1023     1162     PARP.  
FT      CONFLICT        331      337     KGHSLD -> ORPLVA (IN REF. 1).  
FT      CONFLICT        357      361     NEKHP -> IQAS (IN REF. 1).  
FT      CONFLICT        966      966     Q -> P (IN REF. 8).  
SQ      SEQUENCE        1166 AA; 126917 MW; 4CB8BBD97CEF704 CRC64;  
  
Query Match              4.2%; Score 370; DB 1; Length 1166;  
Best Local Similarity    25.6%; Pred. No. 5,5e-12;  
Matches 143; Conservative 61; Mismatches 169; Indels 186; Gaps 14;  
  
OY      19 PALKALLLEKCD-----VDERNESGQ--TPMLAEDGNLEIVELIKNGA 62  
           || :|||::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
DB      23 PAARELEACRNQGVREVRKRLLVPEKVNSRDGTGRKSTPLHFAAGFGRKDVVEYLONGA 82  
           ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
OY      63 NCNLEDJDNMTALISAKBGHNHVVELLCGCYNLEHRDGGTALTMMACYKRGRTIVEL 122  
           ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
DB      83 NWQRDDGGGLPLHNACSFGHALEVNNLLRRGADPNARDMWNTTPLHEAIKKGIDVCIV 142  
           ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
OY      123 LLSHGAMPVS-----TGLOX 137  
           |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
DB      143 LLQDGAPPTLRNDGTATLDADPSAKAVLTGEKKDELLSARSNGEEMALLTLPLNV 202  
           |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
OY      138 SVT-----PTMAAGRGHADIVHLLLONGAKVNCSDK-----YG----- 171  
           |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
DB      203 NCHASDGRKSTPLHLAAGYNRVITQLDLHGADVAHKDKGDVLPLHNACS YGHYEVT EL 262  
           |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
OY      172 -----TTPLVMAARKGHLEVCYKHILLAMGADVNOEGANSATILYA----- 211  
           |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
DB      263 LVNHGACVNAMDIMQFPLRHEASAKRVENCSSLISYGPADPTLLNCHNSAIDLAEPTPOL 322
```

```

QY 212 -----VXG-GYTOSVKE----- 222
DB 323 KERLAEPFGHSLQAREADYRIKKHLSLEWVNKPHQTHETALHCAASPYRKROI 382
QY 223 ---ILKRNPNVLTJDKGNTALMIASKEGTEIVODLDAGTYVNIIDBSGDTVLIGAVR 279
DB 383 CCELLRRKANINKEKTEFLEPLTHVASEKANDVEYVHKAQVNALDNLGQSLHRAAY 442
QY 280 GGHVEIVRALLOKYYADIDIRGONKATLVAVKE-----GNATMVNDILQCNP- 327
DB 443 GCHLQFCRLLSYGCSPNITISLOGFTALQNGNENVOQLLOEGISLGNSEADQLLEAKA 502
QY 328 -DTE-----ICT-----KDGE-----TPLIKATKRNIEVEVELLDKAKYSAVDKGDTP 372
DB 503 GDEYEVKKTCTVGSVNCRODEGRQSTPLHPAAGYNVSVVEYLLQHGADVHAKDKGGLVP 562
QY 373 LHTAIRGRSKLAELLRN 391
DB 563 LHNACSYGHVEYVALLVKH 581

RESULT 12
RNA_HUMAN STANDARD; PRT; 741 AA.
AC 005823;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A. AND MUTAGENESIS OF LYS-240 AND LYS-274.
RC TISSUE-Kidney;
RX MEDLINE=93201598; PubMed=7680958;
RA Zhou A., Hassel B.A., Silverman R.H.;
RT "Expression cloning of 2-5A-dependent RNase: a uniquely regulated
RL mediator of interferon action."
RN Cell 72:753-765(1993).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RX MEDLINE=20515260; PubMed=11063255;
RA Zhou A., Nie H., Silverman R.H.;
RT "Analysis and origins of the human and mouse RNase L genes: mediators
RL of interferon action."
RN Mamm. Genome 11:989-992(2000).
[3]
RP CHARACTERIZATION OF RNASEL ACTIVITY.
RX MEDLINE=94245737; PubMed=7514601;
RA Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,
RA Silverman R.H.;
RT "Intrinsic molecular activities of the interferon-induced 2-5A-
RL dependent RNase."
RN J. Biol. Chem. 269:14153-14158(1994).
[4]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX PubMed=11585831;
RA Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA Saleznada T.;
RT "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RL mitochondrial mRNAs stability in interferon alpha-treated H9 cells."
RN J. Biol. Chem. 276:48473-48482(2001).
[5]
RP ERRATUM.
RA Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA Saleznada T.;
RL J. Biol. Chem. 277:13354-13354(2002).
[6]

RP REVIEW.
RX PubMed=9856285;
RA Castelli J., Wood K.A., Youle R.J.;
RT "The 2-5A system in viral infection and apoptosis."
RL Biomed. Pharmacother. 52:386-390(1998).
[7]
RP MUTAGENESIS OF LYS-392.
RX PubMed=9862963;
RA Dong B., Silverman R.H.;
RT "Alternative function of a protein kinase homology domain in 2',
RL 5'-oligoadenylate dependent RNase L."
RN Nucleic Acids Res. 27:439-445(1999).
[8]
RP MUTAGENESIS OF HIS-583; PRO-584; TRP-632; ASP-661; ARG-667 AND
RX HIS-672.
RP PubMed=11333017;
RA Dong B., Niwa M., Walter P., Silverman R.H.;
RT "Basis for regulated RNA cleavage by functional analysis of RNase L
RL and Treip."
RN RNA 7:361-373(2001).
[9]
RP VARIANTS SER-59; PHE-406; GLN-462 AND GLU-541.
RX PubMed=11941539;
RA Roekman A., Ikonen T., Seppälä E.H., Nupponen N., Autio V.,
RA Mononen N., Bailey-Wilson J., Trent J., Carpen J., Matikainen M.P.,
RA Kolvisio P.A., Tammela T.L.J., Kallioniemi O.-P., Schleutker J.;
RT "Gemline alterations of the RNASEL gene, a candidate HPC1 gene at
RL 1q25, in patients and families with prostate cancer."
RN Am. J. Hum. Genet. 70:1299-1304(2002).
[10]
RP FUNCTION: Endoribonuclease, mediator of interferon action, which
CC play a role in mediating resistance to virus infection and
CC apoptosis. Might play a central role in the regulation of mRNA
CC turnover.
CC -1- CATALYTIC ACTIVITY: Cleaves 3' of UpNP dimers, with preference for
CC UU and UA sequences, to sets of discrete products ranging from
CC between 4 and 22 nucleotides in length.
CC -1- COFACTOR: Optimal RNA cleavage rates requires the presence of
CC either manganese or magnesium and ATP.
CC -1- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
CC linked oligoadenylates) the homodimerization and subsequent
CC activation occurs. Inhibited by RNase L inhibitor.
CC -1- SUBUNIT: Monomer (inactive form) or homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC -1- TISSUE SPECIFICITY: Highly expressed in spleen and thymus followed
CC by prostate, testis, uterus, small intestine, colon and peripheral
CC blood leukocytes.
CC -1- INDUCTION: By interferons.
CC -1- DOMAIN: The kinase domain allows the homodimerization.
CC -1- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
CC constitute the 2-5A binding domain.
CC -1- DISEASE: Gemline defects in RNASEL are linked to cancer-
CC susceptibility for HPC1.
CC -1- SIMILARITY: CONTAINS 9 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
EMBL: I10381; AAA18032.1; -
DR HSSP; P80144; 2MTO.
DR Genew; HGNC:10050; RNASEL.
DR MIM; 180435; -
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00063; ank; 8.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00248; ANK; 8.
DR PROSITE; PS50088; ANK_REPEAT; 6.

```



```

OY      | 365 ILDKAKSYSAVDKKKDDPLHLAIHRSKRSLKEILLRNP----- 392
Db       202 LCKRRASITDC---GDLVA-TARRNYDHSILVXKLISHQAKEDFHPAADMKPQSSSHQA 346
OY       393 -KDGRLLRPKNAGETPYNIDCSHQKSILTTOIFGARHLSPTETDGMGLGY 441
          ||||| |: :|| :| :|||
          347 LKDLHRIVRP-MIGLTKFFIDEKYR-----IADTSEGIIYLGF 383

RESULT 13
ANR6_HUMAN ANR6_STANDARD; PRT; 692 AA.
AC Q9YZG4; G9NUZ4; O9UF09;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Ankyrin repeat domain protein 6.
GN ANKRD6 OR KIAA0957.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=99246063; PubMed=10231032;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RL for large proteins in vitro."
RL DNA Res. 6:63-70(1999).
RN [2]
RP SEQUENCE OF 87-692 FROM N.A. (ISOFORM 2).
RC TISSUE=Testis;
RA Poulikka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 140-692 FROM N.A. (ISOFORM 1).
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE
CC CC PRODUCED BY ALTERNATIVE SPLICING.
CC CC -I- SIMILARITY: CONTAINS 8 ANK REPEATS.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL AB023174; BAA76801.1; -
DR DR EMBL AL117504; CAB55968.1; -
DR DR EMBL AL096678; CAB86658.1; -
DR DR HSSP: O00420; IAWC.
DR DR Genevex, HGNC:17280; ANKRD6.
DR DR InterPro; IPRO02110; ANK.
DR DR Pfam; PF00023; ank; 8.
DR DR PRINTS; PRO1415; ANKRYIN.
DR DR SMART; SM00248; ANK; 6.
DR DR PROSITE; PS50088; ANK_REPEAT; 6.
DR DR PROSITE; PSS0297; ANK_REP_REGION; 1.
KM KM Repeat; ANK repeat; Alternative splicing.
FT FT REPEAT 9 ANK 1.
FT FT REPEAT 38 ANK 1.
FT FT REPEAT 41 ANK 2.
FT FT REPEAT 74 ANK 3.
FT FT REPEAT 107 ANK 3.
FT FT REPEAT 136 ANK 3.
FT FT REPEAT 140 ANK 4.
FT FT REPEAT 147 ANK 4.
FT FT REPEAT 169 ANK 5.
FT FT REPEAT 173 ANK 6.
FT FT REPEAT 202 ANK 6.
FT FT REPEAT 206 ANK 6.
FT FT REPEAT 235 ANK 7.
FT FT REPEAT 239 ANK 8.
```



```

CC -i- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
CC constitute the 2-5A binding domain.
CC -i- SIMILARITY: CONTAINS 9 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce
CC or send an email to license@isb-stb.ch).
CC -----
DR EMBL: AF281045; AAG37308.1; -.
DR EMBL: L10382; AAA37117.1; -.
DR HSSP: P42773; 11TB.
DR MGp: MG1:10962272; Rnasel.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR01245; Tyr_pkinase.
DR Pfam: PF00023; ank; 8.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PRO1415; ANKR1IN.
DR SMART: SM00248; ANK; 8.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TYFKC; 1.
DR PROSITE: PSS0088; ANK_REPEAT; 7.
DR PROSITE: PSS0297; ANK_REPEAT; 1.
DR PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
DR HydroLase: Nuclease; RNA-binding; Endonuclease; Zinc-finger;
KW Repeat; ANK repeat.
FT REPEAT 24 53 ANK 1.
FT REPEAT 58 87 ANK 2.
FT REPEAT 91 120 ANK 3.
FT REPEAT 124 153 ANK 4.
FT REPEAT 167 197 ANK 5.
FT REPEAT 201 234 ANK 6.
FT REPEAT 238 268 ANK 7.
FT REPEAT 272 301 ANK 8.
FT REPEAT 303 328 ANK 9.
FT DOMAIN 364 584 PROTEIN KINASE LIKE.
FT DOMAIN 585 735 RIBONUCLEASE.
FT DOMAIN 229 242 2-5A BINDING (P-LOOP) 1.
FT DOMAIN 253 275 2-5A BINDING (P-LOOP) 2.
FT ZN_FING 401 436 C6-TYPE (POTENTIAL).
SQ SEQUENCE 735 AA: 83274 MW: B66324A5B50F711 CRC64;
Query Match 3.9%; Score 348; DB 1; Length 735;
Best Local Similarity 32.5%; Pred No 4e-11;
Matches 111; Conservative 55; Mismatches 112; Indels 64; Gaps 111
QY 73 TALISAKRGHVIHVEILKCGVNIEN-RDMGWTALIMMACYKGRDVEVLLSHGANPS 131
Db 27 SSLIKAVQGDVRYVQQLLEKSGADANACEDTGWGTPPLHNAVOAGRIDVNLILSHGADP- 85
QY 132 VTGDLQSVYPIIMAGRGHADIIVHLLLNLONGAKVNCDSKYGTPPLVMAARGHGLECYKHTL 191
Db 86 -----HRRKNGA-----TPPTINGDGVKYLETL 112
QY 192 AMGADYDOGANSMTALIVYVAGGYTSYKEILKRNPNVNL---TDKD-----GNLAL 241
Db 113 SCGADVNECEDNGEFTAFMEAAERGNVAELRFLFAKGANVNLRRQTTKDKRRLKOGGATL 172
QY 242 MASAKRGHLEIYODLL-DAGTYVINIPDSGSGTYLIGAVRGHV-----EIVRALLQKYADI 296
Db 173 MSAAEGRGHLEIVIRLLNLMKAEVARDMMGNRNLIRTLMLWDCENVEEISIIIOHGVY 232
QY 297 DIRGDNKALTWAVEKGNATMVRDILQ---CNPDETCIKRQDTEPLIKATKRNIEVE 353
Db 233 NVRGEGKTPPLIAAVERKHTGLVQMLSRBSINIDAR--DNEGKTALLIIVDKQLKEIVQ 290
QY 354 LLLDKGAKVSAVDKKGDTPLHIALIRGSRKLAELL---RNP 392
111:111 111:111 111:111 111:111 111:111 111:111 111:111 111:111 111:111 111:111

```

Tue Jul 1 15:13:30 2003

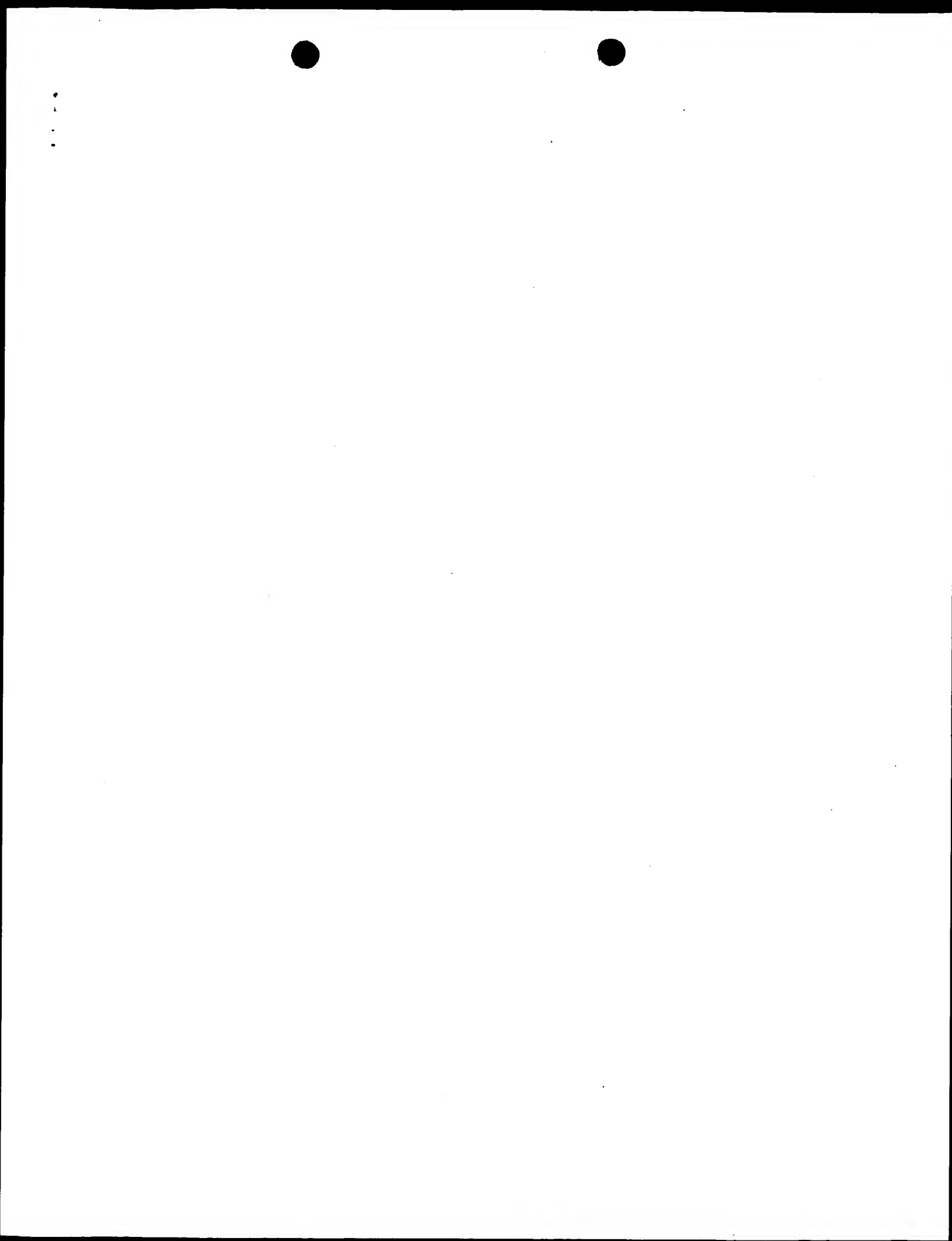
us-10-021-571-4.rsp

Page 19

Db 291 LLEKGA-----DKCDDLWMIARRNDYHLVKLLPYVANP 326

Search completed: July 1, 2003, 14:35:24  
Job time : 38 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 14:34:05 ; Search time 43 Seconds  
(without alignments)  
3834.198 Million cell updates/sec

Title: US-10-021-571-4

Sequence: 1 MSVLISQSVINVEENIPALKECKVDENRGGQPLMTAAAGNLEIKYKLN  
LHAASSESTGFEERESIL 1715

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Database: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5205.5	58.6	1031	2 T43458	hypothetical prote
2	1832.5	20.6	1398	2 T21884	hypothetical prote
3	1255	14.1	246	2 T46392	hypothetical prote
4	555.5	6.3	1881	1 S4HUK	ankyrin 1, erythro
5	552.5	6.2	1856	2 B35049	ankyrin 1, erythro
6	552.5	6.2	1880	2 A35049	ankyrin 1, erythro
7	544	6.1	426	2 AE2149	hypothetical prote
8	539	6.1	3924	2 S37431	ankyrin 2, neurona
9	530	6.0	1961	2 T42716	ankyrin 3, splice
10	529.5	6.0	1940	2 T42715	ankyrin 3, splice
11	525	5.9	1862	2 T49502	ankyrin 3, splice
12	523	5.9	1943	2 T42713	ankyrin 3, splice
13	521	5.9	4377	2 A55575	ankyrin 3, splice
14	516	5.8	1765	2 T42714	ankyrin 3, splice
15	506	5.7	1848	2 S37771	ankyrin 3, splice
16	475.5	5.4	1549	2 T13940	ankyrin, erythrocy
17	442	5.0	2039	2 T15347	ankyrin - fruit fl
18	415	4.7	2584	2 T24158	ankyrin-related un
19	415	4.7	2606	2 T24157	hypothetical prote
20	409.5	4.6	1423	2 T42691	hypothetical prote
21	409.5	4.6	1423	1 T37275	death-associated p
22	387.5	4.4	1062	2 T30255	inversin - mouse
23	375.5	4.2	323	2 B47169	ankyrin-like repa
24	373.5	4.2	557	2 T46507	hypothetical prote
25	373.5	4.2	1062	2 T14151	inv protein - mous
26	349.5	3.9	397	2 T46445	hypothetical prote
27	348	3.9	679	2 B45771	2-5A-dependent RNA
28	343	3.9	741	2 A45771	2-5A-dependent RNA
29	336.5	3.8	991	2 T25412	hypothetical prote

30	329	3.7	1188	2 T19552	hypothetical prote
31	328	3.7	1401	2 S11527	alpha-latrotoxin p
32	325	3.7	934	1 H71274	probable ankyrin
33	320.5	3.6	441	2 S74626	erythroid ankyrin
34	305	3.4	606	2 AC2508	hypothetical prote
35	300.5	3.4	1411	2 S30355	alpha-latrotoxin
36	300.5	3.4	2004	2 D88948	protein ZK1005.1
37	300	3.4	1058	2 D82654	ankyrin-like prote
38	293.5	3.3	627	2 E86190	hypothetical prote
39	288.5	3.2	662	2 E84725	ankyrin-like prote
40	288	3.2	1184	2 T00253	gene Anknz prote
41	286	3.2	517	2 T48283	ankyrin-like prote
42	285	3.2	368	2 T18184	ankyrin repeat pro
43	285	3.2	633	2 T27499	hypothetical prote
44	284	3.2	1083	2 S48460	probable membrane
45	280	3.2	1001	2 S30385	G9a protein - huma

## ALIGNMENTS

## RESULT 1

T43458

hypothetical protein DKFp434F0621.1 - human

C.Species: Homo sapiens (man)

C.Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000

C.Accession: T43458

R.Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A.Reference number: Z22517

A.Accession: T43458

A.Status: Preliminary

A.Molecule type: mRNA

A.Residues: 1-1031 &lt;AAA&gt;

A.Cross-references: EMBL:AL133620

A.Experimental source: adult testis; clone DKFp434F0621

C.Genetics:

A.Note: DKFp434F0621.1

Query Match	Similarity	Score	DB 2;	Length	1031;
Best Local	Similarity	98.6%;	Pred. No. 4.5e-287;		
Matches 1008;	Conservative	2;	Mismatches 0;	Indels 13;	Gaps 1;
QY	1	MSVLISQSVINVEENIPALKECKVDENRGGQPLMTAAAGNLEIKYKLN	60		
DB	1	MSVLISQSVINVEENIPALKECKVDENRGGQPLMTAAAGNLEIKYKLN	60		
QY	61	GANCNLEEDLDMWTFALISASKEGHHVIEELKCGVNLHRDMGWTALMAACYKGRDYY	120		
DB	61	GANCNLEEDLDMWTFALISASKEGHHVIEELKCGVNLHRDMGWTALMAACYKGRDYY	120		
QY	121	ELLISHGANSVYTGLOSYVPIITAAAGRGADIVHLLONGAKYNSDKKGTPLVVAAR	180		
DB	121	ELLISHGANSVYTGLOSYVPIITAAAGRGADIVHLLONGAKYNSDKKGTPLVVAAR	180		
QY	121	ELLISHGANSVYTGLOSYVPIITAAAGRGADIVHLLONGAKYNSDKKGTPLVVAAR	180		
DB	121	ELLISHGANSVYTGLOSYVPIITAAAGRGADIVHLLONGAKYNSDKKGTPLVVAAR	180		
QY	181	KGHLECYVHLLAMGADVDOEGANSMTALIVAAGKGYOSKEILKRPNVNLDDKDGNTA	240		
DB	181	KGHLECYVHLLAMGADVDOEGANSMTALIVAAGKGYOSKEILKRPNVNLDDKDGNTA	240		
QY	181	KGHLECYVHLLAMGADVDOEGANSMTALIVAAGKGYOSKEILKRPNVNLDDKDGNTA	240		
DB	181	KGHLECYVHLLAMGADVDOEGANSMTALIVAAGKGYOSKEILKRPNVNLDDKDGNTA	240		
QY	241	LMIASKEGHEIVODLLDAGTYVNIIPRSGDTVLIGAVRGHVEIVRALOKYADIDIRG	300		
DB	241	LMIASKEGHEIVODLLDAGTYVNIIPRSGDTVLIGAVRGHVEIVRALOKYADIDIRG	300		
QY	241	LMIASKEGHEIVODLLDAGTYVNIIPRSGDTVLIGAVRGHVEIVRALOKYADIDIRG	300		
DB	241	LMIASKEGHEIVODLLDAGTYVNIIPRSGDTVLIGAVRGHVEIVRALOKYADIDIRG	300		
QY	301	QDNKTALYWAVERKNATVWRDILQCNPDTEICTKDGTEPLIKATKMNIEVVELLDKGA	360		
DB	301	QDNKTALYWAVERKNATVWRDILQCNPDTEICTKDGTEPLIKATKMNIEVVELLDKGA	360		
QY	301	QDNKTALYWAVERKNATVWRDILQCNPDTEICTKDGTEPLIKATKMNIEVVELLDKGA	360		
DB	301	QDNKTALYWAVERKNATVWRDILQCNPDTEICTKDGTEPLIKATKMNIEVVELLDKGA	360		
QY	361	KVSAVDDKGDPLPILAIRGSRKLAELLRNPKGRLLYPNKAAGEPVINDSHOSIL	420		
DB	361	KVSAVDDKGDPLPILAIRGSRKLAELLRNPKGRLLYPNKAAGEPVINDSHOSIL	420		
QY	361	KVSAVDDKGDPLPILAIRGSRKLAELLRNPKGRLLYPNKAAGEPVINDSHOSIL	420		
DB	361	KVSAVDDKGDPLPILAIRGSRKLAELLRNPKGRLLYPNKAAGEPVINDSHOSIL	420		
QY	421	TOIFGARHLSPETEDGDMLGVDLYSSALADILSEPTMQPICYGLVQOMSGKSFLLKL	480		
DB	421	TOIFGARHLSPETEDGDMLGVDLYSSALADILSEPTMQPICYGLVQOMSGKSFLLKL	480		
QY	421	TOIFGARHLSPETEDGDMLGVDLYSSALADILSEPTMQPICYGLVQOMSGKSFLLKL	480		
DB	421	TOIFGARHLSPETEDGDMLGVDLYSSALADILSEPTMQPICYGLVQOMSGKSFLLKL	480		

Db 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPICVGLYAQWGSKSFLLKTL 480  
 QY 481 EDEMTFAGQOIEPFPQSWLIVFTLLCGGLGILFFVYHPNUGIAVSTLALTYIF 540  
 Db 481 EDEMTFAGQOIEPFPQSWLIVFTLLCGGLGILFFVYHPNUGIAVSTLALTYIF 540  
 QY 541 FVIVFGGRREGESNNMAMVSTRLARHIGLELLKLMFNPPPELPOTKALPVRLF 600  
 Db 541 FVIVFGGRREGESNNMAMVSTRLARHIGLELLKLMFNPPPELPOTKALPVRLF 600  
 QY 601 TDYNLSSVGGESTLAEMIATLSDACEREFGLATRLFRVETEDTQKKKKTKCLPS 660  
 Db 601 TDYNLSSVGGESTLAEMIATLSDACEREFGLATRLFRVETEDTQKKKKTKCLPS 660  
 QY 661 FVIFETIGCTISGTTLLAIFRVDPKHLTVNNAVLISAVGLAFVNLCTRWQVYDSSL 720  
 Db 661 FVIFETIGCTISGTTLLAIFRVDPKHLTVNNAVLISAVGLAFVNLCTRWQVYDSSL 720  
 QY 721 NSQRRRLHNAASKLHKLSEGMKVLKCEVELMAMAKTIDSFQONQRLVYIIDGLDAC 780  
 Db 721 NSQRRRLHNAASKLHKLSEGMKVLKCEVELMAMAKTIDSFQONQRLVYIIDGLDAC 780  
 QY 781 BODKVLQMLDLYRVLFSGPPLAFASDPHIIKALINONLSVLDNSINGHDYMRIVH 840  
 Db 781 BODKVLQMLDLYRVLFSGPPLAFASDPHIIKALINONLSVLDNSINGHDYMRIVH 840  
 QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDPTGIEDADRRVSONSLGEMTKLSKLTALN 900  
 Db 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDPTGIEDADRRVSONSLGEMTKLSKLTALN 900  
 QY 901 RRDYRRROMQRTIRQMSFDLTLLVTEDMFSDISIPQMRRLNIVSVTGRLLRANQIS 960  
 Db 901 RRDYRRROMQRTIRQMSFDLTLLVTEDMFSDISIPQMRRLNIVSVTGRLLRANQIS 960  
 QY 961 FMDWDLASMINTEQWPRYSMTLIXLEETEGIPQMLKTYER-----IS 1007  
 Db 961 FMDWDLASMINTEQWPRYSMTLIXLEETEGIPQMLKTYER-----IS 1007  
 QY 1008 KNI 1010  
 Db 1021 KSV 1023  
 Db 1021 KSV 1023  
 RESULT 2  
 T21884  
 hypothetical protein F36H1.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T21884  
 J. Kershaw, J.  
 Submitted to the EMBL Data Library, January 1996  
 Reference number: 219482  
 A:Accession: T21884  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1398 <WII>  
 A:Cross-references: EMBL:Z68760; PIDN:CAA92996.1; GSPDB:GN00022; CESP:F36H1.2  
 A:Experimental source: clone F36H1  
 C:Genetics:  
 A:Gene: CESP:F36H1.2  
 A:Map position: 4  
 A:Introns: 38/3; 60/3; 152/3; 191/3; 257/2; 290/3; 323/3; 379/1; 439/1; 485/2; 552/2  
 Query Match 20.6%; Score 1832.5; DB 2; Length 1398;  
 Best Local Similarity 30.9%; Pred. No. 1.8e-95;  
 Matches 468; Conservative 280; Mismatches 483; Indels 285; Gaps 40;  
 QY 34 RNECGQPLMAAEOGNLEIVKELIKNGANCNLEIDLN--WTALISAKSEGHVHVEEL 91  
 Db 58 RNVGESLITVAVRSGNTAVAKQLQALDPDA-IDETDNEMGSALLNAHGHVDPVRLLI 116  
 QY 92 KGVNLHRDMGCGWTALMACYKGRDVEVLLISGANPSVTGLQYSVYPIIWAAGRGA 151  
 Db 1086 -----PLV-----EKIDAVVNLIRKID-IPSNRLDSTLDRFYOL 1119

Db 117 DNGASVDPQDPLMGMSPLMAAVYKNHLDVVDLVN----- 150  
 QY 152 DIYHLLONGAKNCSDKGTTPLVAAARKGHLECKHLLANGADYDQGANSTALTYA 211  
 Db 151 -----AKVNADKREGSTALVIAARKGHLPVVOLLNNGAEVADAGMSTALMA 200  
 QY 212 VKGQTSYKELIKRPNVNLTDKDNFTALMAISKREGHTEIVODLLDAGTYNIDPRSD 271  
 Db 201 TRGNFQVYVLLITREPNVAVDQNGTLTALGMAARGVADICESLNSAFNQCDFRN 260  
 QY 272 TVLIGAVRGHVEIVBALLQKVAIDIRODNKTALYMAVEKNATVBDIQCNDPEI 331  
 Db 261 WILTSAVRGNMAIVRMILDKFADINCOSEKRTPLHAIIDKSFNDIAVILLEKRPNLEL 320  
 QY 332 CTYDGEPLKATKMNIEVELLDKGAQVSVYDKGPPPLHAIARGSRKIAELLN 391  
 Db 321 KNDGEPPLRAKCRVHLCYLMFSGAALVANDCGNALHLALARSRLTQALSLN 380  
 QY 392 PKDGLYRPNKAGETPYNIDCSHOKSILQIQRANHLSPTETDGM--LGDYLYSAL 448  
 Db 381 PDSRLRYRNKLGTPYSIDLSNPQILPLJFG-----PIADKMDTAMGYDVSNTL 435  
 QY 449 ADLSEPTMQPICVGLYAQWGSKSFLLKLEDEKKTAGQOIEPL-FQSWLIVFTL 507  
 Db 436 ADIVCEPSLSLPLTIGIYAKWGSKALAKLEAHMSSESDMLDGVSLVSFALEFALF 495  
 QY 508 LCGGGLAFVYHPNUGIA--VLSFLALTYIF-----IYIFGGRREGESNNMAMV 560  
 Db 496 LFGMSLITPTMLAISNSTAYLISVFLIFITFCSLIVYVYGDGK---NWTSD 552  
 QY 561 LSTRARHIGYLELLKLMFNPPPELPOTKALPVRLFTDYNLSSVGGESTLAEMIA 620  
 Db 553 IANFARVFSRIRLVNLTLLHAP--MSEDSASMPVFLFADYHRLSSIGGEQALAKIYA 611  
 QY 621 TLSDACREREFGLATRLFRVETEDTQKKKKTKCLPSFV-----ILFIIIGCTISGI 675  
 Db 612 TLFEAAHTHGIVPVRIFCCMKPPYRGISLRHRCGHVPVILLIYAVVLLIAQVFTV 671  
 QY 676 TLLAIFRVDPKHLTVNNAVLISAVGLAFVNLCTRWQVYDSSLNSQRRRLHNAASKLH 735  
 Db 672 WILS--DROPNNN--LFLAIAFLCGFVMAIYPLALIMSWTVVPRRVAANRANH 726  
 QY 736 KLSSEGMKVLKCEVELMAMAKTIDSFQONQRLVYIIDGLDACEDQVYQMLDLYRV 795  
 Db 727 KLRREGIMQKIQLEVDLADIMIRSLDAFTSRLVYVVDGLDNCDEBRYOTIDLELL 786  
 QY 796 FS--KGPFLAFASDPHIIKALINONLSVLDNSINGHDYMRIVHLPVFLNSRGLSN 852  
 Db 787 FSARKHRPFTITIAVDPHIVASINNMHSALSGETLGHDIYKNTIISMFYILHNSALQ 846  
 QY 853 ARKFLVSAIANGDVPSCDPTGIEDADRRVSONSLGEMTKLSKLTALNRDPT 905  
 Db 847 LQSKL-----RKRESMAEW-----KEFKKQDFFYGSHTL 878  
 QY 906 -----RRQORITTRQMS-----FDLTLLVTEDMFSDISIPQMRRLNIVSVTGRLLR 955  
 Db 879 READGRSRRKSTVPMNNSNSVYGRMNNDGILLEDYFSSMNPFRAMRIYVNAITLGRKLR 938  
 QY 956 ANQISFMDWDLASMINTEQWPRYSMTLIXLEETEGIPQMLKTYERISKINPTTMD 1015  
 Db 939 AFETIDSEWMSLGHVWSLLEQWPSRMCMLDLRALEVIN--NQLLSVYVQLQDHPAQQD 996  
 QY 1016 VEPLLEIDGDIRNEFEVLSR--TPVIVARDVYKFLPCTVNLDPKLEI-----A 1064  
 Db 997 ---LMQDRNLLENFEGDLDSKGLPSAERLTGVGHKFFVCTSLDLYLKRILREKSGLY 1053  
 QY 1065 DVRAARQJISIGLIAYPPLPLHGGPPRABGYSQPPVSSYFNGPFAAGVAVPOPHSS 1124  
 Db 1054 DIFA--QVSAGAMAI-----PPNARLLFSDTLTMSIDT----- 1085  
 QY 1125 YSGMTGCPHPFYNNRSGPARGVVLLNSLVNAYCEKIKQIEGLDQSMPLPQYCTTIKA 1184  
 Db 1086 -----PLV-----EKIDAVVNLIRKID-IPSNRLDSTLDRFYOL 1119





```

Db      94 VAEVNYGANGVANAQOKGFTPLVMAAQNHELVAFLENGANGVNALEDGFTPLVALAQ 153
      81 EGHVHIVBELKCGV-----NLEHDMGWTALMA 111
      154 OGHENVVAILINVTGKGVRLPALHIAARNDDRTAAVLQNDPDPVLSKGTGFTPLHIA 213
      112 CKKGRDVEVVELLSGANGSVT-----GLQY-----SV 139
      214 AYEENLVAAQOLLNNGASVNETPONGITPLHIASRGNVIMVRLLDREGAOIETKDEL 273
      140 YPIIWAAGRGHADIYHLLONGAKVNCSDKYGTPLVAAKRGHECYHLLAMGADV 198
      274 PHLHCAARNGHVRISIELLDHGAPIQAKTKNGLSPHMAAGDHLDCVRLLDQDAEID 333
      199 -----QEGA-----NSMTALIVAKGITYOSKYEILKR 226
      334 ITLDHLTPLVAAHGHHRVAKVLLDKGAKPNSRALNGFTPLHACKNHVRVEMELLKT 393
      227 NPNVNLTDKDGNTALMIASKEGTEIVODLDAGTYVNIPIRSGDTVLIGAVRGHVEIV 286
      394 GASIDAVTESGLTPHVASFMGHLPIYKLNLRGASPNVSNKYETPLHMAAKRGHTEVA 453
      287 RALLOKYADIDIRGDNKNTALYWAVERGNATVNDIIOCPDTEICTRGDETPLIKATKM 346
      454 KYLLQNKAKVNAKADQDTPHCAIRIGHNMVKKLENNAPNLATAGTAPLHIAARE 513
      347 RIVEVVELLDGAVSAVNDKGDTPHIAIRGRBKRLAELLNPKGRGLLRNKG 406
      514 GHVEVVLALLEKASQACCTKGFPLHAAKYGRVAELLLLE-RDA---HMAAGK 567
      407 ---TPYNDICSHOKSILTOIFGARHLSPTETDGMAGY-DLYSSALADILSEPTQOPIC 462
      568 NGLTPLHVAHNNNDIYKLLRPGSGSPHAPMN---GYTPLHIAKON-----QYVA 618
      463 VGLVYWGSGSKFLLKLEDEKTPAGQOIEPLFQSWLVYFLLLCGG---LGLLPA 518
      619 RSLLYGSGSANAESV-----OGVTPHLH-----LAAQEGHAEVALLLS 656
      519 FTVHNLGIIVSLFLALIIYFIYIFGGRREGESMWAWVLSRLRHIGYLELLK- 577
      657 KOANGNLGKSGLTPLHLV-----AOEG-----HVPVADVLKH 690
      578 -----LMFVNPPELPEQTTKALPVRLF---TDYNRLSVG-----GETSLA 616
      691 GVMVADATPMGYTPLHVASHGNIKLVKFLQHQADVNAKTKLGSPHQAQOCHDIY 750
      617 EMI-----AIIISDACEKEGFLA-TRLFVFKEDTIOGKKKKKTCCLPSFV 662
      751 TLLLNKASPNVSSDGTPLAIARLGIYSTVTLKVV-QDET-----Sfv 796
      663 IFLFIIGCIISITLAIFRVDPKHL-----TVNAVLSIASVVG---LAFVLCRTWMQ 714
      797 L-----VSDKRMSPEVTDIL-DVSDDEEELISKAERK-----832
      715 VLDSLLNSQRKRLHNAASKLRLKSEGFMYLCKEVELMARMAKTIDSTQNOTRLVYI 774
      833 --DSRDVDEKEELLDFVKLDQVESPAPRIPC-----AMPETVVRSEBOQASKEY 884
      775 DGLDACEODKVQMDIVKVLFS--KGFPIAFADP-----HIIK-----814
      885 DEDSLIPSSPATETSDNISPVASPVHTGFVSEFVNDARGSGSRHNGRLRVIPRTCA 944
      815 -----AINONLSVLRDSNINSHDYWRN-IYHLYFVL-NSRGLS 851
      945 APTRTICRLVKOKISTEPPLAEGBGLASRIIALGPTAOPLSPYIVTIPHFASHGK-- 1002
      852 NAKRLVYSAINGDVPCSDTGTGIDADRVQSNGEMTKLGSKTALNRDRYRRRQMO 911
      1003 -DREIVVLRSENGSV-----KREHRSRGESYLDIOLN-GMDEELGSELELEKKRVC 1052
      912 RITROMSFDLTKLVT---EDWFSDISPQIMRLNLTIVSYTGRLLRNQSFPMDRLAS 968
      1053 RIIT--TDPEFLVYVMSRLQD--YDIIGREG-----GSIKSKVLPIVQATPENAVTK 1102
      969 WINTEQMPYRTSMILLYLEETEGIPDQMTK-----TYIERISK-----1008
      1103 RVKLALQ-----AQVPDELTKILGNQATFSPVIVTEPRRRKHPRIQLR 1148
      1009 -NIPPTKDVPLLEIDGDIRNFV-----LSRTPVLVADVVPFLPCTV 1053
      1149 IPLPSTWDPNPRSGESDITSLRLCSYIGCTOAOEMDITGTTKLYVANECAFN---TT 1205
      1054 NLDEKRELIADYRAAREQISIGLAVPPLPLEHGPAPRPSGVSOPVSCSTFNGPPA 1113
      1206 NV--SARFVWLDQDPRTEAVNFATLLYKEL-----TAVPYMAKRV 1243
      1114 GGVYSPGHSYSYSGMTGPHPHYNGSGGAPRV--VLLNSINDANCE-----K 1162
      1244 -----IFAKNDRERLKCCTTDDKVKTYLQEHENVEYAR 1281
      1163 LKOTEGIDQMLPOYCTTIKKNINGRVLAQCNIDELKEMNNPFGDMLFRSTVLEMRN 1222
      1282 SRDIEVLEGMSL-----FALISG-----NLVYKKAQQRSPHFQSPRENRLAM-- 1325
      1223 AESHVPEDEPRFLSESSGAPHPGEAPRASHNE-----LPHTELS-----1263
      1326 -----FVKYVDSREPGSLSFRLKAMKYEYDQHILCHINIMPPCAKSGAEDRR 1376
      1264 -SOTPYTNLSPFELNLTIGDEGAPR-----HSNLSMGOSQTRRP-SLSSLSQD 1311
      1377 RTPPLALRYSLISESTPGSLSTEQEMKMAVISELGLSWELARELOFVEDIN--- 1433
      1312 SSIERSKLTRKVOAEYADAYREYIAQMSQ---EGPGSTTISGRSSPHSTYMGQSSS 1367
      1434 -----RIRENNSNLEQSVALLNLMVIREG-----QANAMENLYTALQSID 1475
      1368 GGSISHNLQDEKGDSEPKDDGKSFMLKRGVDIYSSGVSUNDASPLDPTTEDE-- 1425
      1476 RGEIVNMLESGSGRSRLKPD-----RRHRDPRYSLSPSQMGNYSSL---ODELL 1522
      1426 -----KSDQSGSKL-----LPCKSSERSLSFOTPLKLGSLRQKLPSPDE 1467
      1523 SPASLGALSPPLRADYVMEVAVLDAIPLATEHDTLMESMOMVMSAGLTPSLVTAED 1582
      1468 DE--SGTEESDNTPLDKDKRKAQKVERVPSSEHSAPIRTFKAKEYLSALLDK 1524
      1583 SLSLECSAEDSDAT-----GHEWKLGCALISEBRGE-----LGSLEVEDDTVDS 1628
      1525 KDSS-----DSGVASSSSSPNHSLHNEVADDSQLEKANLIEEDDSHGKRGIPSL 1576
      1629 DATNGLIDLEQEGGSEKRLPGSKRQDQATGAGQ-DSENEVSIVSGHQGQARITHS- 1686
      1577 SGLDPPRIARMSICSEDK-----KSPECSLIASSPEENW 1611
      1687 -----PVSQVTERSDRLQDMADGSIYSYLDQAQGSW 1721

```

RESULT 5  
 B35049  
 ankryn 1, erythrocyte splice form 3 - human  
 N:Alternate names: ankryn 2.1, erythrocyte, ankryn-R  
 N:Contains: ankryn 2.2, erythrocyte  
 C:Species: Homo sapiens (man)  
 C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 10-Jul-1998  
 C:Accession: B35049  
 R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990  
 A:Title: cDNA sequence for human erythrocyte ankryn.  
 A:Reference number: A35049; M01D:90175370; PMID:1689649  
 A:Accession: B35049  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1856 <LAMB>  
 C:Genetics:  
 A:Gene: GDB:ANK1; ANK



```

OY 1426 -----KSDGSKL-----DPGKSSRSLSFOTDLKLGSLRYOKLPSPDE 1467
Db 1523 SPASIGCALSSPLRADQWNEVALIDAIPLATEHDTHLEMSDMQWASAGLTPLSLVRED 1582
OY 1468 DE---SGTEESDNTPLKDDKRAKGEVERVPKSPESHAPRTFTIKAKELYSDALLDK 1524
Db 1583 SLSECKAEDSDAT-----GHEWKLEGALSEPRGPE-----TGLSELEVEDDQVDS 1628
OY 1525 KDS-----DSGVRSESSPNSHNEVAADSQLEKANLIELEDHSHGKGPISL 1576
Db 1639 DATNOLIDLEGEESQSEEEKLPKSKRODDATGAGQ-DSENEVSLVSGHQRQARITHS 1686
OY 1577 SGLDPIIARMSICSEDK-----KSPSECSLIASPEENW 1611
Db 1687 -----PTVSQYTERSDRLQDMADGISTVYLQDAQGSW 1721

RESULT 6
A35049
Ankyrin 1, erythrocyte splice form 2 - human
N.Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N.Contains: ankyrin 2.2, erythrocyte
Species: Homo sapiens (man)
C.Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C.Accession: A35049
R.Lamert, S.; Yu, H.; Prechal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A.Title: cDNA sequence for human erythrocyte ankyrin.
A.Reference number: A35049; M01D:90175370; PMID:1689849
A.Accession: A35049
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-1880 <LAW>
A.Cross-references: GB:M28880
C.Genetics:
A.Gene: GDB:ANK1; ANK
A.Cross-references: GDB:118737; OMIM:182900
A.Map position: 8p11.2-8p11.2
C.Superfamily: ankyrin; ankyrin repeat homology
C.Keywords: alternative splicing; cytoskeleton
F:2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MA>
F:2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA>
F:144-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 552.5; DB 2; Length 1880;
Best Local Similarity 19.6%; Pred. No. 1.2e-22;
Matches 373; Conservative 287; Mismatches 659; Indels 581; Gaps 72;

```

```

OY 81 EGHVHVEELKCGV-----NLEHRDGMGTALMA 111
Db 154 QGHENVAHLINTGTGKYRLPLHLIARNDDTRIAVLONDPNVDLSKGTGFTPLHIA 213
OY 112 CYKGRDVEVELLSHGANSVT-----GLQY-----SV 139
Db 214 AHYENLNVAAQDILNRGSSVFTFQNGITPLHIASRGVNIYVRLDRAQJLETKYKDEL 273
OY 140 YPTIWAAGRHADIYHLLONGAKVNCSDKYGITPLVMAARKHLECYKHLIAGMDV- 198
Db 274 TPLHCAARNGHVRISETLIDHGAPIQAKTKNGLSPTIIMAAQGHLDVCVLLQYDAEID 333
OY 199 -----DEGA-----NSMTALIVAKGQTSVKELR 226
Db 334 ITLDHLTPLVAAACHGHRVAKVLLDKGAPNSRALNGFTPLHIACKKHVNRMELLKT 393
OY 227 NPVNLITDKDNTALMIASKEGTEIVODLDAGTVNIPDRSGDVLIGAVRGHVEIV 286
Db 394 GASIDAVTESGLPLVHASPFGHLPIYKNLQAGASPNSVNVETPLHMAARAQHTVEVA 453
OY 287 RALLQKADIDIGODNKTLIYVAVEKGNATVRDLOCPTEICTKGEFTPLIATM 346
Db 454 KYLLONKAKVNAKAKDQPLHCAARIGHTNMVKLLIENNANPNIATAGHPLHIAARE 513
OY 347 RNFVEVELLDKGAVASVDKGDTPPLHAIIRGRSKLAEILLRNPKDRLLYRPKAGE 406
Db 514 GHVETVALLKEKRAQCAQTKKGTPLHVAAKGKRVAVELLE--RDA-----HRAAK 567
OY 407 ---TPYNIQDSHOKSLITQIFGARHLSPEETGDMIGY-DLYSSALADILSEPTMOPIC 462
Db 568 NGLTPLHVAVHNHNLIDVKKLLPRGSGSPHSPANN--GYPLHIAAKON-----QVEVA 618
OY 463 VGLYAGWGSQSKPLKLEDEMTFAGOOLEPFPQSMILVFTLLCGS---LGLTFA 518
Db 619 RSLQIGGGSANASV-----QGYPLH-----LAAQGHAEVALLS 656
OY 519 FTVHPNLGIIVSLFALTYFEIVYFGGRDEGSMNNAWVLSRLAHIGYELLK 577
Db 657 KQANGNLGNKSGTLPLHLY-----AQCG-----HVPADVLIKH 690
OY 578 -----LAFVNPPELPEQTKALYRFLF---TDYNNLSVG-----GETSLA 616
Db 691 GVMVDAATTRGTYPLVHNSHYGNIKLVKFLQDAQVNAKTKIGSYSLHQAQOQGTDIV 750
OY 617 EMI-----ATLSDACEREGERFLA-TRLPRVETEDTQCKKKKTKCPLSPV 662
Db 751 TLLIKNGASPNVSSDGTPLATAKRLGYISVTDVLYKV-TDET-----SFV 796
OY 663 IFLETTIGLISGTLTIAIFVNDPKHL-----TVNAVLIASVYG---LAFVLCRTWMO 714
Db 797 L-----VSDKHRSFPETVEIL-DVSEDEGELLISFKERR--- 832
OY 715 VLDSLNSQKRRLHNAASKLHKLSSEFMVULCEVELAMAKTKIDSTQONOTRLVLI 774
Db 833 ---DSRDVDEKELDFVPKLDQVESPAIRPIC-----AMETVAINSEDEQASKEY 884
OY 775 DGLDACRQDVLOMLDTRVLFSS--KGFPIAIPASP-----HIIK----- 814
Db 885 DEDSLIPSSPATETSDNISPVATHTGFLVSPVNDARGSMGRHNLGLRVITPRICA 944
OY 815 -----AINQNLNSYLKDSNINSHDYMKN-IVHLVPL-NSRGIS 851
Db 945 APTRTGRLVVKPQKLTSTPLAEGLASRIIALGPTGAQFLSPVIVEIPHFASHRG-- 1002
OY 852 NAKFLVTSATNGDVPCSDTGTGIEODADRVSQNSLCEMTKLGSKALNRDPTRRRQMO 911
Db 1003 -DRELVLVRSSENV-----WKHRSYGSYSDQILN-GMDEGLSLELEKRCVC 1052
OY 912 RTTRQSGFDLTKLVV---EDWFSDISPOTMRRLINIVSGTGLRANQISFMDRLAS 968
Db 1053 RIIT--TDPLFYVIMRLOD--YDIIGPEG-----GSLKSKLVPLVQATFPENAVYK 1102

```

```

QY 969 WINTEQMPYRTRSWLLIYEEETGIPDOMTLK-----TIVERISK----- 1008
Db 1103 RKVLALQ-----AQPPYDELVTYKLGNGATFSPITYVEBRNRKFRPIGLR 1148
QY 1009 -NIPPTKVEPPELEIDGIRNFEYV-----ISSRPVLVARDVYKVEPLCTV 1053
Db 1149 IPLPSPWTDNRDGEEDJTLRLICSVIGTDOAQMEDITGTTKLVANECANF---TT 1205
QY 1054 NLDPKLEIRIADVNARAREOISIGGLAYPPLHGGPPRAPSGYQSPVCSSTFNCPFA 1113
Db 1206 NV--SARFMSDCRRTAFAVFAILLKEL-----TAVPYMAKPV 1243
QY 1114 GGVVSPQPHSSSYSGMTGPOHPFYNRKSGPAPGFV--VLNLSLVDAVCE-----K 1162
Db 1244 -----IPAKMNDPREGRRCYCMDDKVDKTLTQGHENFVAVAR 1281
QY 1163 LKOTEGLDOSMLPOYCTTIKANKINGRYLAOCNIDELKKEKNNMFGWHLFRSTVLEMRN 1222
Db 1282 SRDLEVELEGMSL-----FAELISG-----NLVYPKKAQORSFHFOSFRENRLAM-- 1325
QY 1223 AESHVVEDPPEFLSESSSGPAPHGEPPARASHNE-----LPHTELS----- 1263
Db 1326 -----PYKVDSSREPGSGSLFKRKAKKYEDYQHILCHLNTWPPCAKSGGAEADR 1376
QY 1264 -SQPTPLNFSFEELNLTGLDEGAPR-----HSNLSWOSQTRTP-SLSLSNSOD 1311
Db 1377 RTPPLALRYSLSESTPPGSLSGTEQAEKMAVISEHLGLSMELAELOFVSVDIN--- 1433
QY 1312 SSIISKLTGVQAEYRDAYREYIAOMSOL---EGGPGSTTISGSRSPSTHYMGSSS 1367
Db 1434 -----RIRERNNSLLEQSVALLNLMVIREG-----QANMELTYALDSTD 1475
QY 1368 GGSJHSLNEDEKQKDPKPDGKRKSPLMKRGDIVYSSSGVSTNDASPLDPTTEDE-- 1425
Db 1476 RGEIVNMLEGSGRQSRMLKPD-----RRHTRDYSLSPSQNGVSSL-----QDEL 1522
QY 1426 -----KSDQSGSKL-----LPKKSSEKSLQOTDIKLKSGSLRQKLPSPDE 1467
Db 1523 SPASLGCALSSPLEADYDNEVALTDAIPLAETHDMLKMSQVMSAGLTPSLVLAED 1582
QY 1468 DE--SGTESDNTPILKDKDKRAEKGVERVPKSPHSAEPRTFKAKEYLSDALDK 1524
Db 1583 SSLECSKAEDSDAT-----GHEWKLVEALSEERGPPE-----LGSLELVEDDTVDS 1628
QY 1525 KDSS-----DSGVSSSSSPNHSJLHNEVADDSQLEKANLIELEDDSHSGKRGIPHS 1576
Db 1629 DATMGLIDLLEQEGQRSEKELPGSKRQDQATGAGQ--DSENEVSLVSGHQRGARITHS- 1686
QY 1577 SGLDPIIARMSICSEK-----KSPSECSLIASSPEENW 1611
Db 1687 -----PTVSOVTERSQDRLODMADGSIYSVLDAAOGSW 1721

```

## RESULT 7

AE2149  
 hypothetical protein al12748 [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AE2149  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AE2149  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-426 <KOR>  
 A:Cross-references: GB:BA000019; PIDN:BAW74447.1; PID:g17131841; GSPDB:GNO0179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: al12748

```

Query Match 6.1%; Score 544; DB 2; Length 426;
Best Local Similarity 37.2%; Pred. No. 3.2e-23;
Matches 149; Conservative 64; Mismatches 180; Indels 8; Gaps 6;

QY 9 VINVERENIPALKALEKCKDYDERNEGQFPLTAAGQNLIEYKELKNGANCNL-E 67
Db 8 LKVAKSGDIKIGALLAAGVYDIDCDROGTTALMAANDGTEIYRSLLDGGANVNLAR 67
QY 68 DLDMNTALISASKEGHVHYEELLKCGVNLHNRDMGMATLMACYKGTQDVVELLSHG.127
Db 68 KRYGATLALMAASANOVDYQILLISGCAVNMATNEDGRTALMAALAKGVAVARVYLLANG 127
QY 128 ANPSVYGLOYSYPIIIMAGRGHADVHLLDNGAVNCSDKYGTPLVMAARKHGLECV 187
Db 128 ADVNITDKDDT-ALKLAKRGQAAYVOLLPSGADANCEDEGETLLMLAADSGHGVV 186
QY 188 KHLIAGADVDQEGASMTALIVAIVKGYTQSVKEILKRNPNVNLTDKDGNTALMASKE 247
Db 187 QVLLAGVYDNEQNDGGTALLAAVANGALAKILLDGGADVNNHODGDSALHLYAVE 246
QY 248 GHTEIYQDLIDAGTYNIPDRSGDYVLIGAVRGHVEIYRALLQYADIDIRGQD-NKTA 306
Db 247 GYDVYQVYLLNGANTQINKKLGDTPLVLAALQGHQIYETILL-KYG-ANVHGDMIGETP 304
QY 307 LYVAVERKGNATVVRDILQCNPDTEICTKGETPLIKATMRNIEYVELLDGAVYASND 366
Db 305 LTLAASQGHATYRILLDYGANANIPASOGKTLIATIRNHPGVYQILLAGANVYOD 364
QY 367 KKGDTPLHIAIRGSRKLAELLRNPKDRLLYPRPKAGET 407
Db 365 SVGATALLMAASGYNKVVQILLGEGADTNL--KNRGYGT 402

```

## RESULT 8

S37431  
 ankyrin 2, neuronal long splice form - human  
 N:Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythro  
 N:Contains: ankyrin 2, short form  
 C:Species: Homo sapiens (man)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 13-Aug-1999  
 C:Accession: S37431; A39643; B39643; A40334; A49462; S14559  
 R:Chan, W.  
 submitted to the EMBL Data Library, September 1993  
 A:Reference number: S37431  
 A:Accession: S37431  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-3924 <CHA>  
 A:Cross-references: EMBL:426634; NID:g406287; PIDN:CAA81387.1; PID:g406288  
 R:Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.  
 J. Cell Biol. 114, 241-253, 1991  
 A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal  
 A:Reference number: A39643; MUID:91302466; PMID:1830053  
 A:Accession: A39643  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2077 <ORT>  
 A:Cross-references: GB:X56957  
 A:Accession: B39643  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1443, 3585-3924 <ORT>  
 R:Use, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Wa  
 Genomics 10, 858-866, 1991  
 A:Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.  
 A:Reference number: A40334; MUID:92009921; PMID:1833308  
 A:Accession: A40334  
 A:Molecule type: DNA  
 A:Residues: 463-474, 'PE', 477-495 <TSE>  
 A:Cross-references: GB:M37123; NID:g178647; PIDN:AAA62828.1; PID:g178648  
 R:Chan, W.; Kordell, E.; Bennett, V.





Db 1449 TETSYL--KSH-----LVNEVPLASPDLLSEVSEMKODILKM--TAILT 1489  
 Qy 1184 ANINGRVLAQCNIIDELKEMNNMFGDMLFSTVLEKRNASHVPEDEPRELSE--SSSGP 1242  
 Db 1490 TDVSDKA--GSIKVETLVAAEEEPG-----PFEYER-----VKEDLEVNTILTSGT 1537  
 Qy 1243 APHGPPARASHNELPHTLEISSQTPYTLNFSFEELNLTGLDEGAPRSHNSLWOSQTRTP 1302  
 Db 1538 CTREDSVYOSRSEKGLVE--EEMVIVSDEIEEAR-----QKAPLETTEPCVEYRDLK 1590  
 Qy 1303 SLSLNSQDSSIEISKLTDKQVAYRAYREYIOMSOLEBGPSTTISGRSSPHSTYIM 1362  
 Db 1591 EIKGKVEKSTGLVNTLTDLNT--CVPLPKEDLQTVDDKAKKCEALAVGRSS----- 1642  
 Qy 1363 GOSSGSGSIHNLBOEKDSEP-----KPDGKRSFLMKRGVDIVDSSGVSTN 1412  
 Db 1643 -----EKEKDIIPDETOSTOKQHKPISLGIKKPVYRK----- 1674  
 Qy 1413 DASPLDPTTEDEKSDSGSKLPLPKKSSERSLFTQDLKLKSGLRQKLPSEDESGT 1472  
 Db 1675 -----LKEKOKKEKGLQASAKELKGSSESLGE-----DGLAPEPLPYKATS-- 1722  
 Qy 1473 ESDNTPLLKDDKRAKGEKVERPKSPE-----HSAEPTFTIKAKETLSDALDKDS 1527  
 Db 1723 -----PLTEETPTIGSTIDKVKALQKRYEEDQKRSKLPJR--VKGEDEVPKKTHRPH 1774  
 Qy 1528 SDSGVRSSE-----SSPNHSLHNEVADDSQLEKANLIELEDSDHSGKRGIPHSGLQD 1581  
 Db 1775 AASPLKSERHAPSPSPKTERHSTLSSAKTERHPPS-----PSSKTEKHS 1822  
 Qy 1582 PIARMSICSEDKKSPSECSLIASSPEENWPACOKAYNLNETPSTVTLNNSAPANRANQ 1641  
 Db 1823 PV-----SPSAKTERHSPASSS--SKTEKHSV-----SPSTKTERHSPVSTKTER 1867  
 Qy 1642 NFDEMEGIRETSQYLIRPSSPNPTTQNEULK--SMTHKRSQBSSTYRLSDPP 1694  
 Db 1868 H-----PPVSPSGKTDKRPVSPSGRTEKHPPVSPSGRTEKRLPVSPSGRTDKHOP 1917

## RESULT 9

T42716  
 C:Species: Mus musculus (house mouse)  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
 C:Accession: T42716  
 J:Cell Biol. 130, 313-330, 1995  
 A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene the repeat domain.  
 A:Reference number: 222237; MUID:95340633; PMID:7615634  
 A:Accession: T42716  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1961 <part>  
 A:Cross-references: EMBL:L40632; NID:9710548; PID:9710552; PIDN:AA01607.1  
 A:Experimental source: strain C57BL/6J; kidney  
 A:Genetics:  
 A:Gene: Ank3  
 A:Map position: 10  
 C:Superfamily: ankyrin repeat homology  
 C:Keywords: alternative splicing

Query Match 6.0%; Score 530; DB 2; Length 1961;  
 Best Local Similarity 19.9%; Pred. No. 2,5e-21;  
 Matches 432; Conservative 279; Mismatches 673; Indels 784; Gaps 90;  
 Qy 14 EENIIPALKALECKXDYDERNECGOTPLMAAOGMLEYKELIKNGANCNLEDLDMWT 73  
 Db 66 KEGHEVEVSELLQREANDATKKGNTALHIASLAGAEEVYKVLVTGCAVNAWQSQNGFT 125  
 Qy 74 ALIASKEGHVHIYEELIKGCVNLEHRDMGWTALMACYKGRDVEYELL----- 124  
 Db 126 PLVMAAOGNLEHYVRFLDLNGASQSLATEGCTPLAVLAQGDQVVDVSLLENDTGGKVR 185

Qy 125 -----SHGANSVTGL----- 135  
 Db 186 LPALHTAARKDDTKAAALLLNDNTNADVESKGFPLHTAARHGNTVATLILNRAAYD 245  
 Qy 136 ---QSVYPLIIMAGRGHADIVHLLQNGAKVNSDKYGTTPLVMAARK----- 182  
 Db 246 FTARNIDITPLHASKRGNAWVKKLLIDRGAKIDAKTRDGLTPLHCGARSGHQVEMLLD 305  
 Qy 183 -----HLCYKHLAMGADVQEGANSFTALIVAKGYTOS 219  
 Db 306 RSAPILSKTKNGSLPLMATQGDHLNCOYDILLQHNVPVDVNTDYLTLALHVAHGHGYKV 365  
 Qy 220 VKELIKRNPVNTLDKQNTALMTASKE----- 247  
 Db 366 AKVLLDKKASPNAKALNGFTPLHIAACKNRIRIMELLKHGASIAQVTSGLTPHVAAF 425  
 Qy 248 -GHEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQRYADIDIRGODNKA 306  
 Db 426 MGHVNIYSOLMHHGASPNFTTNVGETALHMAARSQAEEVRYLYODGAQVFAKAKDDQTP 485  
 Qy 307 LYNAVEKGNATMTRDILQCNPDTEICTKGTEPLIATKMRNIEVEYELLDDGAKSAVD 366  
 Db 486 LHISARLGKADIVQOLLQOGASPMNAATTSQTPPLHAREGHEDVAFTLDRGASLSTT 545  
 Qy 367 KKGDTPLHAIIRGRSRKIAEELLR-----NPKDRL----- 398  
 Db 546 KKGFTPLHVAAKGKLEVASLLOKASAPDAKSGSLTPLHVAANHYNOKVALLILDGA 605  
 Qy 399 -YRPNAKGETPYNIDC--SHOKSILTOI--FGAR-----HLSPTETDGMLOY 441  
 Db 606 SPHAAKNGYTPPLHIAAKNQMDIATSLLEYGADANAVTROGIAVHLAAQEGHYDVEL 665  
 Qy 442 DLYSSALADLSEPTQTP-----PLCYGLIAXQM 469  
 Db 666 LLSRMANVNLNNSGTLPLHAAQEDRVNVAEVLVNOGAHVDAQTRMGYTPPLHYGCH--Y 723  
 Qy 470 GSGK--SEFLK--KLEDEKKT-----FAGQOIEPLFOFSMLIVLTLGCG-----LG 514  
 Db 724 GNIKIYVFLQHSAAKNKTKNGYTLALHQAQO-----GHRHITN 763  
 Qy 515 LLEFAVYHPN-----LGIAVSLSFALLIYFPIV-- 544  
 Db 764 VLLQNNASPELTIVNGNTALAIARRLGISYVDTLKYVEEIMTWTITTEKHKNMVPETM 823  
 Qy 545 -----YFGGRREGSMMWAVLSRLARHIGYELIKL-- 578  
 Db 824 NEVLMSDEYRRKASAPKELSDGEYISDEEDGKTFWFKIPKVO-----EVLKSD 875  
 Qy 579 -----MFVNPPELPEQTTKALP-----VRLETFD-----YRLSSVGE 612  
 Db 876 AITGDIDKYIGPDOLKLEGDSDLPAGYVGFSLGASASLRSRSTYLTKRSSIARS 935  
 Qy 613 TSLAEMIALTSDACEREFGLATRLFRVERKTEDTQGGKKKKTKCCLPISVFIPLIIGCII 672  
 Db 936 MMEIEELLVP-----SKEOHLTFTRER-----DSDSLRHYSWADT-----LDWNVLVS 978  
 Qy 673 SGI--TLIAIFRYDPRLHLYNNAVLSIASVGLAFVL-----NRTWQVDSLINSQRKL 727  
 Db 979 SPVHSGFLVSFWADAGSGWRG-----SRHGGRIITIPRCKTAPTRICRLV--KRHFL 1031  
 Qy 728 HN-----AASKLHLKLSSEG-----FMKVLCEVLELMARMAKTIDSFTQ 765  
 Db 1032 ANPPRVVEGGLASRLVENGRPAGQPLGVYIIEPHFGSMRGKERLYLVBRENGETWKE 1091  
 Qy 766 NQ-----TRLVITIDLACEDQDKVQLMDIYRV--LFSKGPPIAFASDPHIITKAINQ 818  
 Db 1092 HQPDSKNEDLAEILNGMD--EELDSPEELGETGRICRIITK--DFQYFA-----VYSRIKO 1143  
 Qy 819 NLNSVLRDSNINCHDMRNINIVHLYPVLNBSRGISNAKPLV--TSATNGDVPSCDTGTGIO-- 875  
 Db 1144 ESNQIGPEEGT-----LSSTTVPLVQASFPBGALTKRIRVGLAQ 1183



QY 876 ---EDARVSNSIGSMWTKGSLTALNRQTY--RRROMQRTTROMSPDLTKL--V 927  
 Db 1184 PVPEEYTKKI-----LGNKATSPITYVPRRRKRKPKPT--MTIPVPPSGEGV 1231  
 QY 928 TEFWFSDISPQTMRELLNIIVSVTGRLLRANOISFNMDRLASWMLTEOMP-----Y 978  
 Db 1232 SNGYKGDAVP-NLRKLSITGGTSP-----AOMEDITGTPPLTFIKDCVSF 1276  
 QY 979 RTS-----WU-----IYLEETBEPDQMTLKT-YERISKNIPTTDFVEPLIDDIR 1027  
 Db 1277 TTINVSARFWLADCHQVL---ETVGLASOLYRELICVPMAFVFAKTNDF---VESLR 1330  
 QY 1028 -----NFEVFLSSRTPVIVARDVYFELPCTVMIDPKLR---EIIDVR 1067  
 Db 1331 CFECMTDDRVDKTEHQENFEVANSK-DIEVLBEGKPIYDVCYGNLAPLTKGQGLVFNFY 1389  
 QY 1068 AARQOISIGLAVPLPLHSGPPAPSGY-----SQPPSCVSSIFNGKPPRAGVYSPQH 1122  
 Db 1390 SFKEN-----RLPPSIKIRDTSOEP--CGRLSF----- 1415  
 QY 1123 SSYSGMTGPQHPFYNRSGPAPGPVVLNSLVANDVACEKTKOLEGDSQMLPQYCTTIK 1182  
 Db 1416 -----LKEKTKTGLPQTVAV----- 1430  
 QY 1183 KANINGRYLAQCINIDELKKEKNMFGMHLFRSTVYLEMNAESHVVPEDPRETSESSGP 1242  
 Db 1431 -CNINILPLPHKKAKEKADRRQS-----FASIALTKRRYS-----YLTEPSSMP 1471  
 QY 1243 APHGEPARR-----ASHNELPHTELTSQTPYTLNFPFELITGLDEGAPRHSNLS- 1293  
 Db 1472 Q---SPEERDIRAIVADHGLSWTELARE---LNFVSVEINOIREVN---PNSLISQ 1521  
 QY 1294 -----W---OSQTRRTPSLISNSQDSSLEISKLDKVOAER-----RDAYREYIAQ 1337  
 Db 1522 SFMLKKRWVRDGNATTDALTSLVTKINRIDYTLLEGPITFDGINISGTRSFADENNVE 1581  
 QY 1338 MSOLEGGRG-----STTISGRSPHSTYYMGSSSGSGLHS--N 1374  
 Db 1582 HDPPVDGHPSPVYELETPMGLYWTPPNPQODHDSLSSISSPRTBRLSDGLVPSQGN 1641  
 QY 1375 LEOBEGK-----KDS---EPKPPDGKRSFLMKRGVDIYSSSGVSTNDASPLDPTIED 1424  
 Db 1642 IEHPTGGPPVYTAEDTSLIEDSKMD-----SVTVDPADPLD--VDES 1682  
 QY 1425 EKSPQSSSKL-----LPKSSSRSS--LFQTLKLKGSGLRQKPLSPDDESGTE--- 1473  
 Db 1683 QLKLCQSECAQCMASVPGLPNDRQAEPLPQTRKVMSEEQEGSGPDEEVTEDKV 1742  
 QY 1474 -----ESDNPPLLKDKDRKAEKVERVP-----KSPESAE-PIRTFIK 1512  
 Db 1743 KSLFEDIQLEVEAEEMTEPDQGMILNRQVGAELAMSLAQOMQETBPGSLSPAQA--- 1799  
 QY 1513 AKETLSALLDKKQSSSGVRS-----ESSPNHSLNEVAVDSQLEKANLIE 1560  
 Db 1800 --RRLTGGLDLRLDDSDQARDSITSYLTGEPKLEANGMHTA--EVLPEA---KAKPYF 1852  
 QY 1561 LEDSHSGKRGIT-----PHSLSGLDOPILIAMSSIOEDKKSSECS--LIASPEENW 1611  
 Db 1853 PESONDIGKOSIKENLKPETHGCGRTTEPV---SPLTAYOKSLEETSKVLEDAPKCV 1908  
 QY 1612 PACOKAYVNLNPTSPVTVLNNNSAPANRANONFDEMEGIRETSOVLIRPSSPNPTTIOE 1671  
 Db 1909 PVGAK--KMTRTAD-----GRARLVQEEEG-----STRSEPK--QGE 1943  
 QY 1672 NUKSMTHK 1679  
 Db 1944 GYKVKTKK 1951

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
C:Accession: T42715  
R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Higgins, A.; Vialamas, M.; Turtzo  
J Cell Biol. 130, 313-350, 1995  
A>Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin ge  
the repeat domain.  
A:Reference number: 222237; MUID:95340633; PMID:7615634  
A:Accession: T42715  
A>Status: preliminary; translated from GR/EMBL/DBD  
A:Molecule type: mRNA  
A:Residues: 1-1940 <PPT>  
A:Cross-references: EMBL:U40632; NID:9710548; PID:9710549; PIDN:AA801604.1  
A:Experimental source: strain C57BL/6J; kidney  
C:Genetics:  
A:Gene: Ank3  
A:Map position: 10  
A:Introns: 834/1  
C:Superfamily: ankryrin; ankryrin repeat homology  
C:Keywords: alternative splicing

Query Match	6.0%	Score 529.5	DB 2	Length 1940
Best Local Similarity	20.1%	Pred. No. 2.6e-21		
Matches 431	Conservative 280	Mismatches 673	Indels 763	Gaps 90

QY	14	EEENIDALKALLECKDVERNEGSGOTPLMTAAGQNIETVKELIKNGANCNIEDNNT	73
DB	66	KEGHVEVSELTÖREANDVAATKKNATLHLSIASLGAQEVVKVLTWGAVNAQSONGFT	125
QY	74	ALLSAREGHVHVEELIKGVNLEHRMGMTAMACVKKGRDVEYELL	124
DB	126	PLTMAAOENHLEVRELLDNGASQSLATEDGETPLAVALAQQGDQVSLLENDTNGKVR	185
QY	125	-----SGANPSVTGL-----	135
DB	186	LPALHTAARKDPTKAAALLQNDTNADVESKSGFTPLHIAAHGNINIVATLLNRAAVD	245
QY	136	----QYSVYPLTMAAGGADIVHLLLDNGAVNCSDKGTTPLVMAARK-----	182
DB	246	FTARNDITPLHASKRGKNNAMVKLLLDGCAIDAKTRDGLTPLHCGARSGHEOVEMLLD	305
QY	183	-----HLECYKHLIAGADVDQEGASMTALVAAKGGYTOS	219
DB	306	RSAPILSKTRKNGLSPLHMAQTGDHLCVQLLDHNRVDDVTYNDITLALHVAAHCHGYV	365
QY	220	VKEILKRNPVNLTDKGNLTALMAKSE-----	247
DB	366	AKVULDKKSPNAKALNGFTPLHICKKNRIRVWELLKHGASIQAVTESGLPIHYAAF	425
QY	248	-GHTEIVODLLDAGIYVTPDRSGCTVYIGAVRGHVEYIRALLÖKTADIDIGQNKTA	306
DB	426	MGHVNIYVQSLMHGASPTNTVNGRTFALHMAARSGQAEVYKYLVDQAOVEAKKDDPT	485
QY	307	LYNAVEKGNATVWRDLOCNPTETCTGDGETPLIKATKKRNIEVVELLDKCAKSAVD	366
DB	486	LHTSARLGKADIVQOOLLQGASPNATTSGITPLHILAREGHEDVAALFLDNGASLSTTT	545
QY	367	KKGDTPLHIAIRGRSKHIAELLR-----NPKDGRLL-----	398
DB	546	KKEFTPLHVAAKYAGKLEVASILLQKSASAPDAAGSGSLPLHVAAHYDNQVALLLLDQGA	605
QY	399	--YRPNKAGETPYNDIC--SHQKSILTQI--FGAR-----HLSTETDGMIGY	441
DB	606	SPHAAKKNGTPTPLHIAKKNMQMDITSLLEYGADANAVTROGIASVHLAAQEGHNDVWSL	665
QY	442	DLYSSALDILSEPTMQP-----PIGVGLYAQM	469
DB	666	LLESRNNAVNSKNSGTLPLHLAAQEDRVNAEVLVNOGAHDAQTKKGTPTPLVGVCH--Y	723
QY	470	GSGK--SFLIK--KLEDEMKT-----FAGQITEPLFQESWMLVETLLLCGG-----LG	514
DB	724	GNKIYVNFLLQSAKAYNAKTKNGYTAGLHQAQO-----GHNHIN	763
QY	515	LLEAFYVHN-----LGIAVSLSFALLIYFFIV-----	543

```

Db 764 VLLQNNASPELTVNGNTALALARRGLYSVVDLKVTEELMTTITTEKRNVPETM 823
Qy 544 --LYFGRRREGSENNAMVLTSLARRHIGYLELLTLK-----MFVNPPLPQOTIK 592
Db 824 NEVLDMSDDDGDCCTFKPKIKVQ-----EVLVASEDAITDITDKYIGFODLKLGD 875
Qy 593 ALP-----VRELFETD-----YNRLSVGGETSLAEMIALSLDACERFGL 633
Db 876 SLPAEGYVGFSLGARSASLSRFSDDRSYTLNFRSSYARDSMMIELLV-----SKQHIL 930
Qy 634 ATPLFRFKTEDVQGGKKKKKTCCLPSVFLFICIIISGT--TLALFRDPDKHLTVN 691
Db 931 FTREF--DSDSLRHSMAADT-----LDNVLVSSPVHSGFLVSPMDARGSGSR 978
Qy 692 AVLIASVYGLAFVL--NCRPMQVLDLSNOKRRLHN-----AASKHLKLS 739
Qy 979 G-----SRHGMGLIIPPRKCTAPRTICLV--KHKLANPPRYVEGLASRLVEMGP 1031
Db 740 EG-----FMKYLKCEVELMARMAKTIDSTONQ-----TRLVIIIGLAGE 781
Qy 1032 AGAQLGVPYVEIPIPHGSMGRERLIVLRSENGETWKEHOFDCKMEDLALLNGMD--E 1089
Qy 782 QDAVLQMDTVRY--LFSKCPFAIFASDPHIITKAINQNLNVLDNSNGHDYRNIV 839
Db 1090 ELDSPEELGTRICRITIK--DEPOYFA-----VSRIKOESNOIGEGGI----- 1133
Qy 840 HLPVFNISRLSNARKFLV--TSATNGDVPCSDTTGIG-----EDADRVSONSLGEMTK 893
Db 1134 -----LSTYTPVLQASPEBALIKRIRVGLQADVPREYTKI-----L 1173
Qy 894 GSTALNRRDTY--RRRQORITTRQMSFDLKL--VTEMFSDISQOTMRRLNIYS 948
Db 1174 GNAKATSPYTVYVPRRRKFKPKIT--MTIPVPPSGEYVNGYKGDATP--NLRLCSTIG 1230
Qy 949 VIGRLIRANGISFNMRLASWMLTEOWP-----YRYS-----WLT-----IYLEE 989
Db 1231 GTPP-----AOWEDITGTPPTFTFKDCVSTTNYNSARFWLADCHQVL--E 1273
Qy 990 TEGIPDOMTLKTI--YERISKNIPTTKDVEPLEIDGDIR-----NFEV 1031
Db 1274 TVGLASQVRELICVYMAKVFYAKTNDP--VESLKCFCMTDRVDKTELOQDNFEE 1330
Qy 1032 FLSSRRPVLYARDKYFLPCTVNLDPKLR--EIIADVARARQISIGGLAVPPLPLHGG 1088
Db 1331 VARSK--DIEYLEGKPIYVDCYGNLADLTGKGOOLVFNYSFKEN----- 1373
Qy 1089 PRAPAGY-----SOPPSVCSSTSFNGPRAGGVSPQPHSSYISGMTGQHPFYNGSGP 1143
Db 1374 --RLPESIKIRDTSQEP--CGRLSF----- 1394
Qy 1144 APGPVVLNLSINLVANDACEKLGLEQDQMLPQYTTIKKANINGRVLACNIDELKEM 1203
Db 1395 -----LKEPKTKGILPQYAV-----CNINILPLHAKRAEKADRDQ 1429
Qy 1204 NNNFGMHLFRSTVLEMRNASHVPEDEPRFLSESSGAPRGEPAR--ASHN 1255
Db 1430 S-----FASIALKRRYS-----VILEPMSPO--SPCFERDIRMAIYADHL 1468
Qy 1256 ELPHTELSQTPYTLNFSFEELNLTGLDGCAPRHSMLS-----W--QOQTRTPSL 1304
Db 1469 GLSWTELEARN--LNFVDEINQIKVEN--PNSLISOSFMLLKWTTRDRKMATDTAL 1521
Qy 1305 SLSNODSSIEISKLTDKVOAEY-----RDAYREYIAOMSOLEGGPG----- 1346
Db 1522 TSVLTKINRDIYTLLEGP1PFDYGNISGTRAFDENNVHDPVDGHPQVELETPMGIX 1581
Qy 1347 -----STTISGSSPHSTYYMGSSGGSIS--NLEDEKG-----KDS--EP 1385
Db 1582 WPPNPFOODDHSIDSIESTSPFRPSRLSDGLVPSQGNIEHPGTGPPVYTAEDSLSDS 1641
Qy 1386 KPDDGKRSFLMKGVDIDYSSSGVSTNDASPLDPIIEDEBKSDQSSKL-----LPCK 1439

```

```

Db 1642 KMD-----SVTVTPADPLD--VDESOLKLCQSECAOCMASVPCIP 1682
Qy 1440 SERS--LFQUDLKSGSLGYOKLPSEDESGTE-----ESDNTPLLKDXK 1485
Db 1683 NGROAEPRLPOTRKRVGSSSEQEKGKSGPDEVEDYKVSIFEDIOLEVEAEEMTDQ 1742
Qy 1486 DRKAEKVERVP-----KSPHSAB--PIRTFKAREYLSDALDKDSSDQSV 1533
Db 1743 GQAMLRVQRAELAMSLAGMONEITPSSLSLESPAQ-----RLTGLDLRLDSDQAR 1797
Qy 1534 SS-----ESSRPHSLHNEVADDSOLEKANLIELEDHSGKRG1-----PH 1574
Db 1798 DSITSVLGEPCIKIANGNHTA--EVIPEA--KAKPYFPESQNDIGKOSIKENLKPKTH 1852
Qy 1575 SLSGLODPIARMSICSEDKKSPSCS--LIASSPEEMNPAQOKAYNINPTPTVTLNN 1632
Db 1853 GCGRTEEPV-----SPLTAYOKSLEETSKIVIEDAKPCVPYGMK--KKTFTTAD----- 1900
Qy 1633 SAPANRANONFDEMEGIRETSOVLIRPSSSPNTTIONENLKSMTHK 1679
Db 1901 -----GKARLNLQEEEG-----STRSEPK--QEGGYKVKTKK 1930

RESULT 11
149502
ankyrin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C:Accession: I49502
R:White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.
Mamm. Genome 3, 281-285, 1992
A>Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory
A:Accession number: I49502; MUID:92345717; PMID:1386265
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1862 <RBS>
A:Cross-references: GB:M4756; MID:g191939; PID:AAA37236.1; PID:g191940
C:Genetics:
C:Superfamily: ankyrin, ankyrin repeat homology
C:Keywords: alternative splicing
F:40-72/Domain: ankyrin repeat homology <AN01>
F:73-105/Domain: ankyrin repeat homology <AN02>
F:106-138/Domain: ankyrin repeat homology <AN03>
F:139-167/Domain: ankyrin repeat homology <AN04>
F:168-200/Domain: ankyrin repeat homology <AN05>
F:201-233/Domain: ankyrin repeat homology <AN06>
F:234-266/Domain: ankyrin repeat homology <AN07>
F:267-299/Domain: ankyrin repeat homology <AN08>
F:300-332/Domain: ankyrin repeat homology <AN09>
F:333-365/Domain: ankyrin repeat homology <AN10>
F:366-398/Domain: ankyrin repeat homology <AN11>
F:399-431/Domain: ankyrin repeat homology <AN12>
F:432-464/Domain: ankyrin repeat homology <AN13>
F:465-497/Domain: ankyrin repeat homology <AN14>
F:498-530/Domain: ankyrin repeat homology <AN15>
F:531-563/Domain: ankyrin repeat homology <AN16>
F:564-596/Domain: ankyrin repeat homology <AN17>
F:597-629/Domain: ankyrin repeat homology <AN18>
F:630-662/Domain: ankyrin repeat homology <AN19>
F:663-695/Domain: ankyrin repeat homology <AN20>
F:696-728/Domain: ankyrin repeat homology <AN21>
F:729-761/Domain: ankyrin repeat homology <AN22>
F:762-794/Domain: ankyrin repeat homology <AN23>

Query Match 5.9%; Score 525; DB 2; Length 1862;
Best Local Similarity 20.1%; Pred. No. 4.4e-21;
Matches 396; Conservative 272; Mismatches 648; Indels 654; Gaps 77;

```

```

Qy 21 LKALLEKKDVERDEGCGTPLMTIAABEGNLEIYKEILKNGANONLEDDWMTLJASAK 80
Db 90 VRELNVNCAVNAOSQKGFTPLYMAOENHLEVKFLENGANONVATEDEGFTPLAVALQ 149

```

QY 81 ECHVHVEELKCGV-----NLEHROGWTALMMA 111  
 Db 150 OCHENVAVHLLVYCKGKVRPLALHIAARNDDRTAAVLLONDPDLVSTGFTPLHIA 209  
 QY 112 CYKGRDYVELLSHANSVY-----GLQY-----SV 139  
 Db 210 AHXYENLVAQILLNNGASVNFPTQNGITPLHIAHSGNVIWVRLLDGAGIETRTDEL 269  
 QY 140 YPIIAAGRGHADIYHLLONGAKVCSKDGTPPLVYAAKRGHIECKHLLAGADVD- 198  
 Db 270 TPLHCAARHNVRISEILLDHGAPLOAKTKNGLSPIHMAAGDHLDVRLLOYNALD 329  
 QY 199 -----NSMTALIYAVKGGYTSQVKEIKR 226  
 Db 330 ITLDHPLPLHVAHCGHHRVAKVLLDKGAKPRSRALNFTPLHACKNHLRVWELLIKT 389  
 QY 227 NPNVNLTDKGNLTALMASKEGHEIYODLLDAGTYVNIIPRSGDVLIGAVRGHVEIV 286  
 Db 390 GASIDAVTESGLTPLHVASFMGHLPIVKNLLORGASPNVSNVKVETPLHMAARAGHTEVA 449  
 QY 287 RALLOKVDIDIRGDNKTALYMAVEKGNATVNRDILQCNPDTEICTDGETPLIKAKM 346  
 Db 450 KYLLONKAKAKAKADDPPLHCAARIGHTGVKLLENGASPNLATTAGHTPLHARE 509  
 QY 347 RNIEVVELLDKGAIVSAVKKDTPPLHAIIRGSRKLAELLRNPKDGRLLYPRNKAGE 406  
 Db 510 GHVDYRALHLEKEASQACMTKGFPLHVAAYKGVRLAELLHEDA-----HPNAGK 563  
 QY 407 ---TPYNDCSHQ-----KSLI-----TOIFGAR----- 427  
 Db 564 NGLTPLHVAHNNNDIYKLLPRGSPHSPAMNGTPLHIAKONQIEVARSLLOYGS 623  
 QY 428 -----HLSPTEDGMDLGYDYSSA-----LADIL 452  
 Db 624 ANAESVQGVTPPLHIAQEGHTEMVALLSKQANGMLKSGITPLHLYSQECHVAVADVL 683  
 QY 453 SE-PTMOPPICVG-----LYAOWSGK-SFLKLEBD-EKTFPAGOOIEPLFOFQSWL- 501  
 Db 684 IKHGTVADATBMGTPLHVASHYGNIKLVKFLQHOADVNAKTKLG--YSPHQAAQOG 741  
 QY 502 -IVFUTLLCGSLGILFAFYVNR-----LGIAVSLFALLIYFFIYIYFGGRRE 551  
 Db 742 HTDIYTLLLKNG-----ASPNVSNNGTTPPLHIAKRLGISTVDLKAVYT----- 786  
 QY 552 GSWMMWAVLSTRLARHIGYLELLKLMFVNPELPEOT-TRKALVRFLEFTDYRLSSVG 610  
 Db 787 ---DETSVVLYSDKHRMSYPTVEIILVSEDEDELVGSKA-----ERRDSRYG 834  
 QY 611 GETSLAEMIALTSDACEREFGLATRLR-----VFKTED-TQKKKKWKTCCLPs 660  
 Db 835 EEKELLDFVPKIDQVVE-----SPAIRIPCVPEYVAVIRSEDEQASKEVEDSLIPS 888  
 QY 661 FYIPLFIIGCIISGITLLAIFRVDPKHLTVANAVLISISVGLAVLNCRTWMOVYDSSL 720  
 Db 889 -----SPATETSDNI-SPVASPVHTGVLVSF-----MVDARG 919  
 QY 721 NSORRKLHNAASKLHLKLSSEGFMKVLCSEVELMAMAKTISFQNOTRLVAVIIDGLDAC 780  
 Db 920 GSMGRSRH-----GLRVYIP-----PRTCAPIRITRILV----- 950  
 QY 781 EODKVLQMDIVRVLESKGPFLAIPASDPHIIKAINONLSVLDNSINGHDYWRN-IV 839  
 Db 951 -----KPKLNTPEPL-----AEEGLASRIITALPTGAQFLSPYIV 987  
 QY 840 HLPVPL-NSRGLSNARKFLVYSATNGVPCSTPTTQDADARRSÖNSLGEKTLGSKTA 898  
 Db 988 ELPHPHSGRG--DRELVLVSENGSV-----WEHKSRRGESYLDILN-GMDEE 1035  
 QY 899 LNRPTVYRRMOQRTITROMSPDLTKLLVT---EDWFDISPTQRRRLINIVSYVGRLLR 955  
 Db 1036 LGSLELEKKRKYCRITIT--TDPLPLFYVIMSRLQOD-YDTIGPEG-----GSLSKLIV 1085

QY 956 ANQISENMDRLASMINLTEQWPRYTSMILLYLEEFGIPDOMTKTYERISKNIPTKOD 1015  
 Db 1086 LVQATFPEENAVTKKVALLO-----AQPVDELVTKL----- 1118  
 QY 1016 VEPLEIDGDIRNEVEFLSKRPVLYARDVKVFLPCTVMDPKRLKITADVNAARQISI 1075  
 Db 1119 -----GNQATF-----SPI-----VVEPERRKFRHPI----- 1141  
 QY 1076 GGLAYPLPLHEGPPRAPSQSPS---VCSSTSFNGFAGVSPQPHSSVYSGMTGP 1132  
 Db 1142 -GLRILPPESTWNPDR-DSEGGDTLSRLCS-----VIGTDQAO-----WEDITGT 1187  
 QY 1133 QHPFYNR-----GSGRPPGVYVNLNL-----NDVACEKL-----QIEGL 1169  
 Db 1188 TKLILNECANFTTNYSARFMLSDCRTEAIVAFATLLYKELTAVPYMAKFIYAKMDA 1247  
 QY 1170 DQMLPOYCTTYK-----ANINGVLAQCNIDELK 1200  
 Db 1248 REGRLRCYCTMDKVDKTELOHENVFVARSRDIEVLEGMPLFAELSG-----NLVPRK 1301  
 QY 1201 KEMNMNGDHLPRSTVLEMRNASHVPEDPFLSESSSGRPPHGEPARASHNE----- 1256  
 Db 1302 KAAQONSFHQSTRENRILAI-----PVKVDSSREROGFLSFKTKMYEDTGH 1350  
 QY 1257 -LPHTELS-----SQPTYLNFSEELNLTGLDEGAR-----HS 1290  
 Db 1351 ILCHLNTMPPTCKSGAEDRRRTLPPLRYSLSESRLGFTSDRREMYAVIREHL 1410  
 QY 1291 NLSMOSQTR-----RPPSLSSUNSOOSI-----RISLTKQVQAEVYD 1329  
 Db 1411 GLSMAELARLOFSVEDINRIRVENPNLSLDOSTALLTLMVDREGNMAK-ENLYTALKN 1469  
 QY 1330 AYREYIKOMSQLEGGPSTTISGRSPHSTYYWGSSGSIHNSLEOKERKDSPEKPD 1389  
 Db 1470 IDRSEIYNM-LFG-----SGROS-----NLRKERH 1495  
 QY 1390 GRKSLMKRGVDIYSS-----SGVSTNDASPLDPTIED 1424  
 Db 1496 GDRSEYSLSPQVNGVSLDDELSPASLOAYALPCLADQYMEVYVDAIPL-NAATED 1554  
 QY 1425 ---EKSPQS---GSKLLPGKSSERSSLFQV-----DIKKGSGLYQKLP-----SD 1466  
 Db 1555 TMLMSDMQVWSAGLTPSLVTAEDSSLECKADSDAIPWKLEGHASHTDQPELGSD 1614  
 QY 1467 EDESGTESDNT-----PLKDDKRAEGVEVERPSPHSAPITFTIKAEYISDALI 1522  
 Db 1615 LVEDDITVDSDATNGADLIGQEGGQRSEKRGQV-SGTQDPTETVSLVSGQORVARTT 1673  
 QY 1523 DKR-----DSSDSGRSSSESPNHSILNEVADDSQLEKRLIELEDSSHGKRGIT- 1572  
 Db 1674 DSPSVRQVLDRSQARTLMDKQGSTAVHPQEAQSSMOE---EYVQGFHSFORRITIO 1729  
 QY 1573 ---PHISGLQDPLIARMSICSEDKKSPSCSILASSPENMPACOKAVN 1619  
 Db 1730 GPEPGALQYEYQVYV--STREHVQKRPETGSPKAGKEPSIMAPESAFS 1776

## RESULT 12

T42713

ankyrin 3, splice form 1 - mouse

C:Species: Mus musculus (house mouse)

C&gt;Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000

C/Accession: T42713

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Elcher, E.M.; Hlgins, A.; Yialamas, M.; Turtzo

J. Cell Biol. 130, 313-330, 1995

A:Title: Ankyr3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge

the repeat domain.

A:Reference number: 222237; PMID:95340633; PMID:7615634

A:Accession: T42713

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1943 &lt;PEP&gt;

A:Cross-references: EMBL:140632; NID:g710548; PID:g710550; PIDN:AA01606.1

A:Experimental source: strain C57BL/6J; kidney  
C:Genetics:  
A:Gene: Ank3  
A:Map position: 10  
A:Introns: 855/1  
C:Function:  
A:Description: supposed to play an important role in the polarized distribution of many  
A:Note: major kidney ankyrin  
C:Superfamily: ankyrin; ankyrin repeat homology  
C:Keywords: alternative splicing

Query Match	5.9%	Score 523;	DB 2;	Length 1943;
Best Local Similarity	20.2%	Pred. No. 6.1e-21;		
Matches 437;	Conservative 270;	Mismatches 660;	Indels 800;	Gaps 90

[illegible]

Db	850	--SDGEBGDAITGDTDKLGPQDLEKJGDSLPAEGYGFSLGARSASLRFSFSDRST	907
QY	647	QCKKWKTKCCLPSPYVIFLPTIGCIIISGTTLLAIRVD- KULY-----VNAVLSI	697
Db	908	INR5STAR---DSMATELWPSKEOHLTFTRFSDSLRHYSAADTLVDVNVISPV	963
QY	698	ASVAGLAFLNCRITMMQVLDLSLNSORRKLHAAS----- KLHKU- 737	
Db	964	HSGFLVSFVNDARG-----GSMRGRHHGMKIIIPRCKTAPRTCTCLVRRHKIA	1011
QY	738	-----KSEG----- FMKLVCEVEMAMAKTIDSFTON	766
Db	1015	NEPPWVEGEGLASRLVEMGPAGAOFIAPYIVEIPHFGSMRGKRELIVLRSENGETWKEH	1074
QY	767	Q-----TRVYIIDGLDACEODKVLOMDTVAV- LFSGPRPIALFASPHIIIAINON	819
Db	1075	QDSDKNEDIAELLNOMD- EELDSEELGTKRICHRIYK- DPPQYFA-----VSRIOKE	1122
QY	820	INSLVRDSNINCHDMRNIVHLPLVFLNSKGLSNARKFLV- T5ATNGDVPCSDPTQIO--	875
Db	1127	SMQIGPEGI----- LSTVPLVQASFPEGALTKRIRVGLQAO	1166
QY	876	--EDADRVSQNSLGEMLTKSGKTALNRDIT- RRQOKRPIITQOMSEDTKLK-----VT	928
Db	1167	VPEETVKKI----- LGKKAFTSPVIVPEPRRRKFKHPIK- MTIPVPPSGEGVS	1214
QY	929	EDWFSDISIPQTRRLINIVSVYGRLLRANQISFNMDRLASVNLTEOMP----- YR	979
Db	1215	NGYKDDARP-NIRLCSITGGSF----- AQMDHITGTPITPLTRKCVASF	1255
QY	960	TS-----WL----- ILYEETEGIPDOMTKTI-YERISKNIPTTKDVEPLLEIDGDIR-	1027
Db	1260	TNVSARFNLADCHQVL----- ETVGLASOLYRELICVPMAKFVFAKTNDP-- VSSLRIC	1311
QY	1028	-----NFEVLSRTPVLVADVAVKPLPCYVINDPKLR- EIIADYRA	1066
Db	1314	FCMTDRDRDKTLEQOENEFVARSK-DIEVLGKRTYDCCGNLAPLTKGGOULFENYS	1377
QY	1069	AREQISIGGLAYPLPLLEHGPPRABSGY----- SOPSPVCSSTPNFGPAGGVSPQPHS	1122
Db	1373	FKEN----- RLPFSIKIRDTQEP-- CGRLSF-	1397
QY	1124	SVYSGMTGPQHPHYTKRGSGPAPGVVNLNSLVNDAVCEKLOIBELDSMLPQYCTTIKK	1186
Db	1388	----- LKEPKTKGLPOTAV-----	1412
QY	1184	ANINGRVLACONIDELAKEMNMNFEDWHLFRSTVLEMNAESHVVEDPREFLSESSGPA	1243
Db	1413	CNLTITTLPAHKHAKKADRRQS----- FASIALRKRTS----- YLTFESMSPQ	1455
QY	1244	PHGEPARR----- ASHNELPTELSSQTPYTLNFSFEEELINTJGLDEGAPRHSNL-	1293
Db	1455	--SPECETDIRMAIVADHLGLSTWELARE----- LNFVDEINOIRVEN-- PNLSLIQS	1504
QY	1294	-----W----- QSOTRRPPLSLSLNSODASIELSKTLDVQAEY----- RDAYREYINQM	1338
Db	1505	FMLKKNVTRBGKNAKTTALTSLVLTAKIRRIDVYTLLEBPJEDYGNISGTRSPADNNVFH	1564
QY	1339	SQLEBGPB----- STTISGRSSPHSTYYVQOSSGSGSIHS-NL	1375
Db	1565	DPVDOHPSFOYLEETPMGLWYTPMPNPQODDHFSIDISIEPFRPRLSDGIVPSQONI	1624
QY	1376	POEGK----- KDS-- EKPDDCKRSFLMKRGVDIDYSSSGVSTNDASPLDPTTEDE	1423
Db	1625	EHPYGGPPVYTAEDTSLDSKMD----- SVYVYTDPADPLD- VBESQ	1665
QY	1426	KSDPSGSKL----- LPKTKSSERS-- LFQOTLKLSGGLRYOKLPDEDESGTE--	1473
Db	1666	LKDLQCSACQCMASVREIPDNGQAEPLRQOTRKVGMSSQOEKRGKSPDEEYTEDKYK	1723
QY	1474	-----ESDNTPLLADKDKRAEGKYERV- KSPESHAB- PIRTFIKA	1513
Db	1726	SLFEDIOLEEVAEEMTDOGOAMINRYORAEIAMSLSLAGONETPPSSLESPQOA-----	1781



Db 1103 LLMGDELDSPPEELAGKRLICRLITKDF-----POYFAVVS--RIKOSNQICPE 1150  
 Qy 951 GRLLRAN-----QISFMMDRLASINLLEQMPYKSWLILYLBETBEPQOMLKYER 1005  
 Db 1151 GGLISSTTVLVQASFEFGALTKRIRVGLQ-----AQPVPDEIKYKILGNK 1196  
 Qy 1006 ISKNIPYTKDVEPLELIDGDIRNEFVLSRTPYLVARDVKVFLPCTVNDPLKRLRIAD 1065  
 Db 1197 AT-----FSPVIVTEPRRRKHKRITWTIP----- 1222  
 Qy 1066 VRAAREOISIGGLAYPPPLPHIEPPRAPSGYSDP--PSVCSSTSFNGPAGVVSPOPHS 1123  
 Db 1223 -----PPSGEGVSNYKGDTPNLRLLCSIT-----GGSP----- 1254  
 Qy 1124 SYSGMTGPQHPFYNRSGSPAPGVVLLN-----SLAVD-----AVCEKIKQIEGLDQM 1173  
 Db 1255 AQMEDITGTT-----PLTIKCCVSEFTTNVSAFELADHQVLEVTGLATOL 1301  
 Qy 1174 -----LPOYCTTIKKANINGRVLAOCNI-----DELKK--EMNMNFGDHLFRSTVLEM 1220  
 Db 1302 YRELICVPYAKFVFAKMDPVESLSRCFMTDDKDKTLLEQENFEE-----VA 1352  
 Qy 1221 RMASHVVPEDPRELSSSSGSPAPHGEPARRA-----SHNELPHT-----ELSSQTPYT 1269  
 Db 1353 RSKDIEVLEKRPITY--DCYGNLAPLTGGOOLVFNFYSPKENRILPFSIKIRDTSOEPCGR 1411  
 Qy 1270 LNFEEELNTLGLDEGAPRHSNLSWOSQTRTPSLSLNSODSSIEISKLTQVQAEYRD 1329  
 Db 1412 LSPFAKEKRTKGLPQTVAVCNLNTILPAHKETES-----DODEIE--KTDRQSHASL 1463  
 Qy 1330 AYRE--YIAOMSOLEGGPGSTTISGRSSPHSTY-----YMGSSSGGSIHNSLE 1376  
 Db 1464 ALRRKRYSYLTPEGMIEKSTGAT--RSLP--TTYSYKPFSTRPYQSWTAPITV----- 1513  
 Qy 1377 QEKGDSEPRPDGKRFELMKRGVIDYSSSGYSTNDASPLDPI-----TEDEKSDQSGS 1432  
 Db 1514 -----PGPAKSGF-----TSLSSSSNTPSASPLKSIWSVSTPSPIKSTLGAS 1556  
 Qy 1433 KLLPGKKSERSLSLFQDLKLGSGLR--YOKLPSEDESGT-----EESDNPPL----- 1480  
 Db 1557 TTSSVKISIDVYASPIRS--LRTMSSPIKTVYOSQSPYNIQVSSGTLARAPANTEAPLKGLA 1615  
 Qy 1481 -----LKDDKRAKAEKVER-----VPRSPF-----HSAPIRTFT--KAEKYLSDAL 1521  
 Db 1616 SNTFFSRTSPYTAGILERSISITWTPRASPKNIMYSSSLPFRSIIITSAPALISSPL 1675  
 Qy 1522 -----LDKSSDSGVSRSSESPNHSL--HNEVA-----DDSOLEKAN--LIE-- 1560  
 Db 1676 KSVVSPYKSRVDYISSAKITMASSLSPPVKQMGHAEVALVNGSISPLKAYSSSTLINCQ 1735  
 Qy 1561 -----LEDDSHSGKRGIPHSLSGLOPITIAHMSICSEDKSPSECSLIASSPENNPACQ 1615  
 Db 1736 KATATLOEKISATNNSVSVSAATIDVEKVFSTTAMPSPILR--SYVSAP----- 1786  
 Qy 1616 KAVNLANFTBSTVTLNNSAPANANONFDEMEGIRETSQVILRPSSS-----PNPTTTON 1670  
 Db 1787 SAFOSLRTPEASALYSLGSSISATTS-----SVTSSIIITVYVSVVAVLEPAPLKL 1839  
 Qy 1671 ENLKSMTKRSQSS--YTRLKSDPPPELHMAAS 1702  
 Db 1840 PDSNSFTKSAALLSPIKITLTETHPOPHSRSTS 1873

# RESULT 14 T42714

C:Species: Mus musculus (house mouse)  
 C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
 C:Accession: T42714  
 R:Peterson, L.L.; Johnson, K.M.; Lu, F.M.; Elcheb, E.M.; Higgins, A.; Yialamas, M.; Turitzo, J.  
 J. Cell Biol. 130, 313-330, 1995  
 A>Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene  
 the repeat domain.

A:Reference number: Z22237; MUID:95340633; PMID:7615634  
 A:Accession: T42714  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1765 <P>  
 A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1  
 A:Experimental source: strain C57BL/6J; kidney  
 C:Genetics:  
 A:Gene: Ank3  
 A:Map position: 10  
 A:Introns: 1587/1  
 C:Superfamily: ankyrin; ankyrin repeat homology  
 C:Keywords: alternative splicing  
 Query Match 5.8%; Score 516; DB 2; Length 1765;  
 Best Local Similarity 19.8%; Pred. No. 1.3e-20;  
 Matches 408; Conservative 253; Mismatches 636; Indels 762; Gaps 83;  
 Qy 14 EENIPALKALLEKCDVDERNBCCGTPIMIAEONGNEIYKELIKNGANCLIEDLDMWT 73  
 Db 66 KEGHVEVVELLDREANVDAAATKGTALHTIASLAGQAEVAVLVYNGANVAASQNGPT 125  
 Qy 74 ALISAKEGHVIHVEELKCGVNLERHDMGWTALMMACYKGRDVEELL----- 124  
 Db 126 PLYMAQENHLEVYREFLNDNGASQSLATEDGFTPLVALQGHQDVYSLLENDTKGKYR 185  
 Qy 125 -----SIGANPSYVGL----- 135  
 Db 186 LPAHTAARKDDTKAALLLNDNTNADVESKSGFTPLHIAHYGINAVATLLNRAAAD 245  
 Qy 136 ---QYSVYPLIIMAGRGADIVHLLONGCAKVNCSOKYGTPLVMAARKG----- 182  
 Db 246 FTARNDITPLHASKRGNAMKLLDRKAKIDATROGTLPLHCGAASGHEQVYEMLLD 305  
 Qy 183 -----HLECYHLLAMGADVQDEGANSMTALYAVKGYTQS 219  
 Db 306 RSAPILSKTKNGLSPLHMAQGDHLNCVQLLQHNVPVDVYNDYTLALHVAHCGHYKV 365  
 Qy 220 VREILKRNPNVLTGKDGTAALMASKE----- 247  
 Db 366 AVYLDKRAKSPAKALNGFTPLHIAKRRIRVMELLKNGASIQAVTESGITPLHVAAF 425  
 Qy 248 -GHTTEIVODLDAGYVNPDPDSGPTVLIGAVRGHVEIVRALLOKYADIDIRGODNTA 306  
 Db 426 MGHVNVYQSLDMHNGASPNPTNVRGELHMAARSOAEVYRVLVDGQAEKAKADDTP 485  
 Qy 307 LYVAVEKGNATVRILOCPTEICTYKDGFTPLIKATKRNIEVEELLIDGAKVSAVD 366  
 Db 486 LHSARLGRADIVQOILDOGASPNAAITSGYTPHLHAAREGHEVAAFLDHGASLSITT 545  
 Qy 367 KKGDPRLHAIAGRSKLAELLRNPKDGRLLYRNNKAGEPYNIDCSHOKSILQITGA 426  
 Db 546 KKGFTPLHVAAYGKLEVASILLQKSSAS--PDAAG-----RSGILPLHVA 588  
 Qy 427 RHLSPTETVDGMLGYDLYSSALADILSEPTQPIYCVGLYAQWGSKGSFLKLLEDEKMT 486  
 Db 589 AH-----YDNQKVAL-----LLDQGASPHAAK----- 612  
 Qy 487 FNGQOIEPLPFSWLVFTLLLLCGGLLFAFTVHPNIGIAVSIETALLIYFTIYTF 546  
 Db 613 -----NGYTPRLHAAKNOMDIATSL-----LEY 636  
 Qy 547 GG-----RREGESWMAWVLSRLRHHGYLELLKLHFNVPPELPBQTTALPVPRLFT 601  
 Db 637 GADANNAVTRQG-----IASVHLAAQREHVDVMSLLSRN-----A 671  
 Qy 602 DYNRLSVGGERTSLAEMITATISDACEBFGFATLRFVFTYE-----DROGKKKW 652  
 Db 672 NVN--LSNNSGLRPL-----HLAAGEDVNVVAEVLVYNGAHVADQTKRGY 714  
 Qy 653 KTKCCLPSTVIVFLFIIGCISGITYLLAIFRVDPKHLVYNAVYISTASVYGLAVYLCRTW 712  
 Db 715 TP-----LHVGHCHYGNIKI-----VNFLLQHSAKV-----NAKT- 743



```

QY 713 MOVLDSLNSQRRRLHNAASKLHKLKSGFMKVLKCEVELMARAKTIDSTQOTRLV 772
Db 744 -----KNGYTAHQAAQGHTHIINV-----LLONNASPNELTVNGNTALAI 785
QY 773 IIDGAGEODKVLQMLDTVRLVFLSGPFIAIFASDPHIIITKAINONLSVLRSGNIGH 832
Db 766 -----ARRLGISVTVTLKVV--TEELMTTTTITERIKM--NPEPMNEVLDMSD----- 831
QY 833 DYMRNIIVHLPVFLNLSRGLSNARK-----PLVTSATNGDPCSPDTGIGEDADRVSQNS 886
Db 832 DEVRKASAPKLSDEYISDEBEGDKTWFKIPKQOEVLRKSED--AITGDTKYLQPOD 889
QY 887 LGEM-----TLGSKTA-LNRDQYRRRQMRITTRQMSFULTKLV----- 927
Db 890 LKELDDSLPAGYVGSLSGASASLSRFSDDSTYLLRSYARDSMITEELVPSKEQH 949
QY 928 ---TEDWFSDISPQTRRLNIVSVTGRLLRANOISFNW--DRLASWINTLEOMPRTSW 982
Db 950 LFTREPDSD--SLRH-----YSWADTLDN-VNLVSS-PVHSGF 985
QY 963 LILYEETRGIPDOMTLTYERISKNIPTKQVPELLEIDGDIRNEVFLSKRPVIVA 1042
Db 966 LVSFWDARG-----GSMK-----GSR-----H 1003
QY 1043 RDVVFELPCTVNLDPKLEIILADVAREQISIGLAVPLPLHEGPPRAPSXSQSPSV 1102
Db 1004 HOKRIITP-----PKCTAPRITICRLYKRHLNAMP--PMWEGEGLASRLVEMGA- 1053
QY 1103 CSSTSFNGPAGVVSPOPHSSYY-----SGMTGPOHPFYNGSGAPGEPV 1149
Db 1054 -----GAOFLGPVIVEIPHFGSMRGKEREILVLESGEETWKEHOFKSNEDLAE----- 1103
QY 1150 LINSLVANDAVCKLQIISGLQSMK-----POYCTIKKANINGVLAQCINDELK 1201
Db 1104 LINGMD-----EELDSPELGTKRICRITTKDPQYFAVVS-----TKO 1143
QY 1202 EMMNMFQDMLFRSTVLEMRNA-----BSHVVED--PRFL-SESSSGPA 1243
Db 1144 EBNQIGPBGILSSTVPLVQASPEGALTRIRVGLAQVPEETVAKIIGNKATFEP 1203
QY 1244 PHGEPARRASHELPH-----ELSS 1264
Db 1204 VTVEERRRRFHKPIMTIPVPPSGEGVNGYKGDATPNLLKLSITGTSFPAQWEDTIG 1263
QY 1265 ORP-----YTLNFS-----FEELNTIGIGDEBAPRHSNLSMOSQTR-----TPS 1303
Db 1264 TTPLETFIKDCVSTTNVSAKREWLADCHQVLETVGL-----ASQLYRELICVPY 1311
QY 1304 LSSL-----NSQDSIEISKLTU-----KVQAEYRDYREYIAQMSQLEGGPGSTT 1349
Db 1312 MAKFWVFAKTNDPVSSLRFCMTDRVDKLTLEQENFEVARS--KQIEVLEGPPIVD 1369
QY 1350 ISGRSSPHSTTYMGSSSG-----SISNLEQE-----KGDSEPRKDDGKSTFLMRG 1399
Db 1370 CYGNLAP-----LTKGGQOLVFNFSFKENRLPLFSIKIRDSQEP--CGRLSFLK-- 1418
QY 1400 DVIDSSSGVSTNDASPLD-----PLTEDEKSD--OSGSKLLPGKK-----SERESL 1446
Db 1419 ---PRTTKGLPOTAVCNLITIPPAKKAERKADROSFSALRKRYSTLTPSMSPOSFC 1475
QY 1447 FQTDKL-----KSGGLRYQKLPSDEDESGTE-----ESDNT-----PLLK-----DDK 1485
Db 1476 EKTDRMAIVADHLGSLWELARELNFVDEINQIRVENPNLSISQSMLLKKNWTRGK 1535
QY 1486 DRKAG-----KVERVP-----KSPH 1502
Db 1536 NATTALTSVLTKIRIDIVTLLEGPIDYGNISGTRSFADENNVFHDVPDQWQONETPSG 1595
QY 1503 SAE--PIRPIKAKEYLSDALLDKXSSDSGVSS-----ESSPHSLHNEVAD 1549
Db 1596 SLESPPQA-----RRLTGGLDRLDSSDQARDISTSYLTGEPKTEANGNHTA--EVIP 1648

```

```

QY 1550 DSQLEKANLELEDSDSHSGKRG-----PHSLSGLODPITARIAMSCSEDKSPSECS- 1601
Db 1649 EA--KAKPYPPSONDIGKOSIKENLKRTHCGGTTEPV-----SPLTAIYKSLSEITSK 1701
QY 1602 -LIASSPENWPAQOAYNINRTPSTVTTLNNSAPARANONFEMEIGIRETSOYLARS 1660
Db 1702 LVIEDAPKPCVPVGMK--KMTRTTAD-----GKARLNLOEBEG-----S 1738
QY 1661 SSPNPTTIQENIKSMTHK 1679
Db 1739 TRSEPK--QGEYKVKTKK 1755

RESULT 15
S37771
ankyrin, erythrocyte - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C:Accession: S37771
R:Birkemeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found
A:Reference number: S37771; M01D:93252825; PMID:8486643
A:Accession: S37771
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1848 <BIR>
A:Cross-references: EMBL:X69063; NID:g311816; PIDD:CAA48801.1; PID:g311817
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:48-80/Domain: ankyrin repeat homology <AN01>
F:81-113/Domain: ankyrin repeat homology <AN02>
F:114-146/Domain: ankyrin repeat homology <AN03>
F:147-175/Domain: ankyrin repeat homology <AN04>
F:176-208/Domain: ankyrin repeat homology <AN05>
F:209-241/Domain: ankyrin repeat homology <AN06>
F:242-274/Domain: ankyrin repeat homology <AN07>
F:275-307/Domain: ankyrin repeat homology <AN08>
F:308-340/Domain: ankyrin repeat homology <AN09>
F:341-373/Domain: ankyrin repeat homology <AN10>
F:374-406/Domain: ankyrin repeat homology <AN11>
F:407-439/Domain: ankyrin repeat homology <AN12>
F:440-472/Domain: ankyrin repeat homology <AN13>
F:473-505/Domain: ankyrin repeat homology <AN14>
F:506-538/Domain: ankyrin repeat homology <AN15>
F:539-571/Domain: ankyrin repeat homology <AN16>
F:572-604/Domain: ankyrin repeat homology <AN17>
F:605-637/Domain: ankyrin repeat homology <AN18>
F:638-670/Domain: ankyrin repeat homology <AN19>
F:671-703/Domain: ankyrin repeat homology <AN20>
F:704-736/Domain: ankyrin repeat homology <AN21>
F:737-769/Domain: ankyrin repeat homology <AN22>
F:770-802/Domain: ankyrin repeat homology <AN23>

Query Match 5.78; Score 506; DB 2; Length 1848;
Best local similarity 19.4%; Pred. No. 5.2e-20;
Matches 378; Conservative 284; Mismatches 653; Indels 634; Gaps 73;

QY 21 LKALLECKVDNEGCGTFLMIAEGNLELYKELINGANCNLEDLNTALISAK 80
Db 98 VRELVTNGAVNNAQSQGFTPLMAOENHLEVKFLLENGANONVATEDGFTPLVALQ 157
QY 81 EGHVHVEELLKGV-----NLEHRDMGWTALMA 111
Db 158 QGHENNVANHLINTGKGVRLPALHTAARNDRTRAAVLLQNDPNPNDVLSGTGFPPLHIA 217
QY 112 CYGRIDVVELLSHGANSVT-----GLQY-----SV 139
Db 218 AHENLVNAQOLLNRGASVNFPTQNGITPLHISRRGNVIMVLLDRGAQIETRTKDEL 277
QY 140 YPIITMAAGRHADIVHLLQNGAKVNCSDKYGTPLVMAARKGHLECYHLLAMGADV 198
Db 278 TPLHCARNGNHVRISLEILDHGAPIOAKTKNGLSPIHMAAQGDHCDVRLLOYNALIDD 337

```



```

QY 199 -----ORGA-----NSMALLVAVGYSVKELTKR 226
Db 338 ITDHLPLRVAHAGHHRVAKVLLDKGAPNSBALNGTPLHACKNNIRWELTKT 397
QY 227 NPNVNLDDKGNALMTASKEGHEIYVODLLDAGTYVNIIPRSGDVLIGAVGHEIY 286
Db 398 GASIDAVTESGLPLHVASFMGHLPIYKNNLQSGASPNVSNVVERPLHMAAAGHTEVA 457
QY 287 RALLQKADIDIKQDNKLTALYMAVEKGNATYVRDILQCPDTEICTKDETPPLIKATM 346
Db 458 KYLLQNKAKANAARAKDQPLHCAARIGHGMVKLEENGASPNLATTAGTPLHMAARE 517
QY 347 RNEIYVELLDKGAYSADVCKGDTPLHAIIRGSRKLEALLRNPDSGLLRPNKAGE 406
Db 518 GHVDPLALLLEKESQCMTKKGFTPLHVAKATGYKRLAELEHDA-----HPNAGK 571
QY 407 ---TPYNIDCSHO-----KSL-----TOIFGAR-----427
Db 572 NGLLPLHVAHNNHNDIVKLLRPGSGPSHPANNGYTPPLHIAKQNOIEVARSLQYGS 631
QY 428 -----HLSPEEDGMIGYDLYSSALADILSEPTWOP-----PIC---462
Db 632 ANAESVGVTPPLHAAQEGHTEMVALLSKQANGNIGKSLTPLHLYSQEGHVPADVL 691
QY 463 -----VGLYAQMSGK-----SFLKKLED-----EMKTFAGQOTEPLFQPSWL- 501
Db 692 IKHGVTYDATRKGTYPLHVAHSHGNKLYKFLQHOADVNAKTKG--YSPHQAQOQ 749
QY 502 -IYFLLLLCGGLGLFAFTVHPN-----LGIAVSLSPALLIYFPIYIFGGRRE 551
Db 750 HTDIVTLTKNG-----ASPNEVSSNGTPLLAIAKRLGISYVDVLYKVI-----794
QY 552 GEGNNAMVYSTLARIIGYLELLKLMFNPPELQOTKALPY--RFLFDYRPLSS 608
Db 795 -----DETSVVLVDKHMASYPETVDELIVS-----EDGETAHISIMDELIVGSKARRS 846
QY 609 --VGETSLAEMIATISDACEREFGLATRLFR-----VEKTED--TQKMKKKKTC 656
Db 847 RQDGEKELDDPKLDQYVE-----SPAIPRIPCVTPPTVYIRSDDQASKEDEDS 900
QY 657 CLPSEVLEFLEICISIGITLLAIFRPDKHLTVNANVLISIASVGLAFVLCRTWQVYL 716
Db 901 LIPS-----SPATERSDNI--SPVASPVHGTGLVSP-----WV 931
QY 717 DILLNSORRKLHNAASKLHKLKSEGMKVLKCEVELMARRAKTIDSTONQRLVYIDG 776
Db 932 DARGGSMGRSRN-----GLRVYIP-----PRCAAPTRITCRLV-----966
QY 777 LDACEODKVLQMLDTRVLFSSKGFPIAIFASDPHIIKAINONLSVLDSNINHDMR 836
Db 967 -----KQOKINTPPL-----AEEGLASRIITALGPTGAOPLS 999
QY 837 N-IVHLEPVL-NSRGLSNARKLTVSATNGDVPSCDTGTIOEDADRRVSONSLGEMTKLG 894
Db 1000 PVIIEIPHFASHGRG--DRELVLRSSESV-----WKHKSRIGESYLDQILN-G 1047
QY 895 SKTALNRDQYRRRQORQTTROMSFDLKLVT---EDWFSDISPQIMRRLINIVSVTG 951
Db 1048 MDEELISLELEKERRCRIT--TDPELVFVIMSRLOCQ-YDTIGPEG-----GSLRS 1097
QY 952 RLFRANOISFNMRRLASMINLLEOWMPYRISWMLLYLEETEGIPDOMTLTIYERISKNP 1011
Db 1098 KIVPLVQATPEPAVYNKKYKLAQ-----AOPVDELVTYKLL-----1134
QY 1012 TTADVPELEIDGIDIRNEFEVLSSRPPVLAVDKVFLPCTVNLDEKREIILADVARE 1071
Db 1135 -----GNQATE-----SPI-----YVEPRRRKFRHPI-----1157
QY 1072 QISIGGLAVPPLPLHSGPPRABSGYISQPS--VCSSTSPNGFPAGVAVSPDHSYSYG 1128
Db 1158 -----GLRIPLPSPMTDNR-DSGEGDITSLRLCS-----VIGTDQAOQ-----WED 1199

```

```

QY 1129 MTGQHPHFYNR-----GSGPARGPVYLNLSL--NDAVCEKTL-----Q 1165
Db 1200 ITGTTKLIYANECANFTTVNSARFMLSQCPRAEAVHERATLYKELFVAPYAKVPIPAK 1259
QY 1166 IEGLOSMPLPOYCTTIK-----ANINGRVLAQONI 1196
Db 1260 MDAAREGRILCYCMDDKVDKTLQEOHENEVEARSDIEVELEGMPLFAELSG-----NL 1313
QY 1197 DELKEMANNFGDMLFRSTVLEMRNASHVVEDEPRELSESSSGPARGEPARASHNE 1256
Db 1314 VPKKAAOORSEFHOFOSFRENRLAI-----PVKVRDSSRBPGEFLSPFRKTKYE 1362
QY 1257 -----LPHTELS-----SOTPTIMFSPBELTGLIDGAPR-----1288
Db 1363 DPHHILCHNITMPTCKGSAEDRRRTYPLTLRYSLISESLGFTSDTDREVMRAVI 1422
QY 1289 --HSLNWSQOTR-----RTPSLSLNSQDSSI-----ELSKLTDKVOA 1325
Db 1423 REHLGSLMELARELOFVEDINRIKRVENPNLDOSTALTLTWYDREGENAKM-ENLYT 1481
QY 1326 EYRDVREYIAQMSOLEGGPGSTTISGRSSPHSTYVMGQSSGSIHSLNLEQEKDSEP 1385
Db 1482 ALRNIIDRSEIVMLE-----VSGROS-----RNLKP 1507
QY 1386 KPDDGRKSPLMKRGVYIDYSS-----SGVSTNDASPLDPTIEDEKSDQSGSKLLPCK 1439
Db 1508 ERHGGREYSLSPSOVNGYSSLODELLSPASLOVALP--SPICADQYMEVYITAIPLAA 1566
QY 1440 SERSSLFQYDLKLGSGLRQKLPSEDE--SGTEESDNPPLKDKDKRAEKGVBRV 1496
Db 1567 TEHDPTLEKSDMOVMSAGITPLSLVYAEBSLSCSAEBSDALP-----EMKLEGAHSED 1620
QY 1497 PKSPEHSAEPT---RTFIKAEVYSDALLDK-----DSSDSGVRSSES 1537
Db 1621 TQGPGLSODLVEDDYVSDATNGLADLLGQORVHARITDPSVYQVLDRSQARFLDWDK 1680
QY 1538 SPNHSILHNEVADDSOLEKANLLELDDSHSGKRG-----PHLSGLQDPIIARMSIC 1590
Db 1681 QOSTAVHPEAOSSMOE---EVQGPSPHFQRRITTTIQQBPBALQOYEQVLY--STR 1733
QY 1591 SEDKKSPECSLJASSPEENMPACOKAYN 1619
Db 1734 EHVQGRPPTGSPKAGKESPLMAPESAPS 1762

```

Search completed: July 1, 2003, 14:43:12  
 Job time : 79 secs

